

FIGURE 1

GGGGCTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCCCTTCAGCCCTGTAAATTGGACATCTGCTGCTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATAAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGAGTTATGCATTG
CTACCATTTATGTTGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTTGTACTTGAATACTGAGTTGTTAGGACTTCTATTGTGGCAAACCT
CCAGAAAACAACCCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTTGACACCACCTTCTACCAAATGCAGCCCCAAATCCATGGCAAACAGTC
TTCTGGATCAGACTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGGACTGATTAGAACAGAAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTCCTCTT
GGTTTTCTGACTTACATTGATTTGAGAAAATTCTTACGGGTGGAAGCCAATTACA
TGGATTAACCCCTATGACACTGCACCTGCCATTAAACATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAAGGATAAAATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTCTGAAATTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAACACCTATGCCTATACTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFIRLLLVIWCGVSAALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFLTYIRDQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGCAGTCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGTGGCTCGGGAGTGGAAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTTGGGTGGCTTCTCATGCGCCAATTGTTAAAGACTATGAGATACTGAGTCAGTATGTTGACAG
GTGATCTCTCCGTGACGTTGCATTTCTGCACCATGTTGAGCTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCGTTATTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCTGG
TTTCATGGTGCCTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCATAAACAGA
CTGCTTTCTGTCTTATGGCTGACCTTATGTATTCTCTGGAAACTAGGAGATCCCTTCC
CATTCTCAGCCAAAACATGGATCTTATCCATAGAACAGCTCATGCCGTTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTGGATTGGTGTCACTGCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCAGTGTGCAAACCATGGATATGAT
CATAAAGAAAAAGAAAGGATGGCAATGGCACGGAGAACAAATGTTCCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAGTGTGTTACCACTCAGCATCAGGAAGTGAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAAGCAGGAGCTTCTGGAAACAGCTGATCT
ATATGCTACCAAGGAGAATAGAAATACTCCAAACCTCAAGGGAAATATTTAATTCTGGTT
ACTTTTCTCTATTACTGTGTTGGAAATTTCATGGCTACCATCAATATTGTTTGATCGAGTT
GGGAAAACGGATCCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTGGTCCCAACACATTCTCATTCTGGAAATACTCGTCACATCCATCAGAGGAT
TGCTGATCACTTACCAAGTTCTTATGCATCTAGCAGTAAGTCTCCAATGTCATTGCTCTG
CTATTAGCACAGATAATGGCATGTACTTGTCTCCTCTGTGCTGATCCGAATGAGTATGCCTT
AGAATACCGCACCATAATCACTGAAGTCCTGGAGAACTGCAGTTCACTCTATCACCGTTGGTTG
ATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCCTCTATTGGCTACAAACAGGCACCA
GAGAACAAATGGCACTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAAATTAA
GATATAAGAGGGGGAAAATGGAACCAGGGCTGACATTATAAACAAACAAAATGCTATGGTAGC
ATTTCACCTTCATAGCATACTCCTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCGTGTGGATATGAGGCTGG
TGTAGAGGCCAGAGGCCAGAACAAACTAAAGGTGAAAATACACTGGAACCTGGCAAGACATGT
CTATGGTAGCTGAGCCAAACAGCTAGGATTCCGTTTAAGGTTACATGGAAAGGTTATAGCTTG
CCTTGAGATTGACTCATTAACACTAGAGACTGTAACAAAAAAAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAACGTTGGCCGCAACTTGTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLVCILLILVFMPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGD
FPILSPKGILSIEQLISRGVIGVTIMALLSGFGAVNCPTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALEELS
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLL
AQIMGMYFVSSVLLIRMSMPLYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLA
HQAPEKOMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCACTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
TCCATCTGGACCACGAGGCTCTGGTCCAAGGCCTTTGCGTGCAGAAGAGCTTCCATCCAGGT
GTCATGAGAATTATGGGGATCACCCCTGTGAGCAAAAGGCACCCAGCAGCTGAATTTCACAG
AAGCTAAGGGAGGCTGTAGGCTGCTGGACTAAGTTGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGCTTGAAACTTGAGCTATGGCTGGGATGGAGATGGATTGTCATCTCTAG
GATTAGCCAAACCCCAAGTGTGGAAAATGGGGTGGGTGTCTGATTGAAAGGTTCCAGTGA
GCCGACAGTTGAGCCTATTGTTACAACACTCATCTGATACTTGACTAACTCGTGCATTCCAGAA
ATTATCACCAACAAAGATCCCATTCAACACTCAAACACTGCAACACAAACACAAGAATTATGT
CAGTGACAGTACCTACTCGGTGCATCCCCACTCTACAATACCTGCCCTACTACTACTCCTC
CTGCTCCAGCTTCACTTCTATTCCACGGGAAAAAAATTGATTGTCACAGAAGTTTATG
GAAACTAGCACCATGTCTACAGAAACTGAACCAATTGTTGAAAATAAGCAGCATTCAAGAATGA
AGCTGCTGGTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGCTCTCCTCTGGTGTG
CAGCTGGCTGGATTGCTATGTCAAAAGGTATGTGAAGGCCTTCCCTTACAAACAAGAAT
CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCATGATAGCAACCTAA
TGAGGAATCAAAGAAAATGATAAAAACCCAGAAGAGTCAAGAGTCAAGCAGCAGCAGTGC
GATGCCCTGGAAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CATGCTCTTACCCGCCAGCTGGGAAATCAAAGGCCAAAGAACCAAAGAACAGAAATCCA
CCCTGGTTCTAACTGGAATCAGCTCAGGACTGCCATGGACTATGGAGTGCACCAAAGAGAAT
GCCCTCTCCTTATGTAACCTGTCTGGATCCTATCCTCCTACCTCAAAGCTCCCACGGCT
TTCTAGGCTGGCTATGCTTAATAATATCCACTGGGAGAAAGGAGTTTGCAAAGTGCAGGAC
CTAAACATCTCATCAGTATCCAGTGGTAAAAGGCCCTGGCTGTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTACTGATTCCGAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAACACGTATCCACCTGACATGTCCTCTGAGCCCGTAAGAGCAAAGAAT
GGCAGAAAAGTTAGGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTGTAAA
GCTAAAATAAGAAATAGAACAGGCTGAGGATACGACAGTACACTGTCAGCAGGACTGTAAAC
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGAATCACTGTTAGAACACACACA
CTTACTTTCTGGCTCTACCACTGCTGATATTTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAAATCTTATAAAATTCTATTTCTGAGTTAGAACATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAGTGCTGTGCAAGGTATTACACTGTAATTGAATTCTCTGAAACACATTGCTGAATAGCT
AACGCTATCTGGGAAGCTATTTCTGAGTTGATATTCTAGCTTATCTACTTCAAACCTAAT
TTTATTTGCTGAGACTATCTTATTCTCTAATATGGCAACCATTATAACCTTAAATT
TATTATAACACACAAAGTACATTGTTACCTCTATATACCAAGCACATTAAAAGTGC
ATTAACAAATGTATCACTAGCCCTCTTCAACAAAGAAGGACTGAGAGATGCAGAAATATT
TGTGACAAAAATTAAAGCATTAGAAAACCTT

FIGURE 6

MARCFSLVLLTSIWTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNCGKNGVGVLIWKPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKA AFKNEAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFFPTNKNQQKEMIETKVVKEEKANDSNPNEESKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCGCGGGCCGCCACCGCGCCGCTCCGCATCTGCACCCGCAGCCCCGG
GGCCTCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGAGCGCAACTCGGTCCAGTCGGGGCGG
CGGCTGCGGGCGCAGAGCGGAGGATGCAGCGCTTGGGGCCACCCCTGCTGCTGCTGCTGGCGG
CGGCGGTCCCCACGGCCCCCGGCCGCTCCGACGGCACCTCGGTCCAGTCAAGCCGGCCCG
GCTCTCAGCTACCCGAGGAGGCCACCTCAATGAGATGTTCCGAGGGTTGAGGAACTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGATGAACCTGGCAAACTACCTCCAGTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAATAATACCATCCATGTGCACCGAGAAATTCAAGATAACCAACAAACCAGAC
TGGACAAATGGTCTTTCAGAGACAGTTACACATCTGTGGGAGACAAGAAGGAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTGCCAGCTCCAGTAC
ACCTGCCAGCCATGCCGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTCTGGGGTCACTGCACAAAATGCCACCAGGGGAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCAGCCGGGCTGTGCTGCTTCAGAGAGGGCTGCTGTCCCTGTGTC
ACACCCCTGCCGTGGAGGGGAGCTTGCATGACCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCTGTGGAGCCTGGACCGATGCCCTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGCCAGCTTCATGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCCTGGGGAGCCGTGCG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGTAGATGTGCAA
TAGAAATAGCTAATTATTCCCCAGGTGTGCTTAGGCGTGGCTGACCCCTTCTCCTA
CATCTCTCCAGTAAGTTCCCTCTGGCTTGACAGCATGAGGTGTGCTGCAATTGTCAGCT
CCCCCAGGCTTCTCCAGGCTCACAGTCTGGTCTGGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTGCCACCCCTGTCCAGATTATGGCTGCTTGCCTCTACCAGGCTGAGACAGCCG
TTTGTCTACATGGCTTGATAATTGTTGAGGGAGGAGATGAAACAATGTGGAGTCTCCCT
TGATTGGTTGGGAAATGTGGAGAAGAGTGCCTGCTTGCAAAACATCAACCTGGCAAAATG
CAACAAATGAATTTCACGCAGTTCTTCATGGCATAGGTAAGCTGTGCCTCAGCTGTTG
AGATGAAATGTTCTGTCACCTGCATTACATGTGTTATTCA
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCAGCACAGCCTGGG
AGGGGGCATTGTTCTCCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTGTACTCTAACGCTCAGTGTCT
CTCCACTACCCACACCAGCCTGGGCCACAAAAGTGCTCCAAAAGGAAGGGAGATGGGAT
TTTCTGAGGCATGCACATCTGGATTAGGTCAAACTAATTCTCACATCCCTCTAAAGTAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGAGCCGTCTTAATGAAGACAATGAT
ATTGACACTGTCCTCTTGGCAGTTGCATTAGTAACTTGAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAAATTGTAAGGGCGAGGATTATAAATGAAATTG
TGTGTGAAACATGGTTGAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG
TTTCAGGTGTCATGGACTGTTGCCACCATGTATTCACCAGAGTCTTAAAGTTAAAGTGTCA
CATGATTGATAAGCATGCTTCTTGAGTTAAATTATGTATAAACATAAGTGCAATTAGAA
ATCAAGCATAAAACTCAACTGCAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLNLPPSYHNETNTDTKVGNNTIHVREIHKITNNQTGQMVFSE
TVITSGDDEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQRMCTRSECCGDQLCVWGHC
TKMATRGSGNTICDNQRDCQPGLCCAFQRGLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGIELLPREVPEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACGCGTGGGGCTGTGAGAAAGTCCAATAAATACATCATGCAACCCAC
GGCCCACCTTGTGAACTCCTCGCCAGGCTGATGTGCGTCTTCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGTCTGGGGCTCTCTGGACCCTT
AACTGGGTACTGCCCTGGCCAATGCGCTCGCTGGAGCCTTGCCCTTACTGGCCTT
CCACAAGCCCCAGGACATCCTACCTTCCCCTTAATCTCGCTCATCCGACACTCCGTTACC
ACACTGGGTCATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCCGGGTATCTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGCCGCTGCATCATGTGCTGTT
CAAGTGCTGCCTTGGTGTCTGGAAAAATTATCAAGTCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTCTCAGCCAAAATGCGTCATGCTACTCATGCGAAAC
ATTGTCAGGGTGGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTGGAAAGCTGCTGGT
GGTCGGAGGCGTGGGGCTCTGCTTCTTCTCCGGTCGCATCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGCTAT
GTCATGCCCAGCGGCTCTCAGCGTTTCCGCATGTGTGACACGCTTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGCCCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCCAGGACTGCACCCCACCCCAACGTCCAGCCATCCAACCTCACTCGCCTTACAGGT
CTCCATTTGTGGTAAAAAAAGTTTAGGCCAGGCGCGTGGCTACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCCGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGT
AAACCTCCGTCTTATTAAAAATACAAAAATTAGCGAGAGTGGTGGCATGCACCTGTCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACTCCAACCTGGTGACAGACTCTGTCTCCAAAACAAAACAAAACAAA
AAGATTTTTATTAAAGATATTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLOIYGVGLFWTL
NWVLALGQCVLA~~G~~A~~F~~ASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKF~~I~~KFLRNAYIMIAIYGKNFCVSAKNAFMILLMRN
IVR~~V~~VVLDKVTD~~L~~LLFFGKLLVVGGVGVL~~S~~FFFFSGRIPGLGKDFKSPHLNYWLPIMTSILGAY
VIASGFFSVFGMCVD~~T~~LFLCFLEDLERNN~~G~~SLDRPY~~M~~SKSLLKILGKKNEAPPDNKKRK~~K~~

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGGCCCGGCCGCGCCGGCGCCCAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
TCCTGCTCAGCTGCGCTCCTGCCTCTGCGGCTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTCACGTTCTTCTCTTCCCTGGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCGGCGTGGAGAGTCAGCTACAAGCTGCCCTGGGTGTGAGGAGGGGCC
GGGATCCCCACCGTCTGCAGGGCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCCCTCTTCTTCTTACCCCTGCTCATGCTCTGCGTGAAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTTGGTTCTTAAGTCCGTACCTGGTGGCCTC
ACCGTGGGTGCCTCTACATCCCTGACGGCTCCTCACCAACATCTGGTCTACTTCGGCGTGTGG
CTCCTTCCCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTGCACACTCTGGAACCGAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCAGTCCCGTGCCTGGTACGCAGGCCCTTCTTCACTCTCCTC
TTCTACTTGCTGTCGATCGCGCCGTGGCGCTGATGTTCATGTTACTACACTGAGCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCTAACCTCACCTCTGTGCTGCCTGGTACGCAGGCCCTTCTTCACTCTC
CCAAGGTCCAGGACGCCAACGCCAACCTGGGCTGCTGCAGGCCCTGGTACCCCTCACCATG
TTTGTCACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGGCAACGAGACAGTTGTCGAGGCCGAGGGCTATGAGACCCAGTGGTGGGATGCCGAGCATTG
TGGGCTCATCATCTTCTCTGTGCAACCTCTTCATCAGTCTGCCTCAGACCACCGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCACCTATGCTAGGCCACACAGCAGCAGCAGCAG
GGCAGCCTGTGAGGCCGGCCTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCTTCTC
ACTCTGCCTGGTGTGGCTACTGCACGTACATGATGACGCTACCAACTGGTACAAGCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCTCTGCAGCAACCGCAGACTCAGCTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCTGCCACCTGGTGCCTCTGGCTCGGTGACAGCCAACCTGCCCTC
CCCACACCAATCAGGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGC
CTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGGCTGCCTCTCCCTCCCTCCCTGGTGTGGCCATACTCAGCATCTGGATGAA
AGGGCTCCCTGTCCCTCAGGCTCCACGGGAGCAGGGCTGCTGGAGAGAGCAGGGAACTCCCACACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTGGTACAGTCCCCCAGGGACCCCTGCCCTTCC
GACTTCGTGCCTTACTGAGTCTAAGACTTTCTAATAACAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLGKAE
ECDSRAWYAGLFFFLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCFLFISLRSSDHRQVNLSLMQTEECPPMLDATQQQQQOVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWKICASWAGLLLLWTLVAPLLLNRND
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCACCCGTTAAGGTGTCTTCTCTTAGGGATGGTGAGGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGGCCAGAGCTCCCATGCTTCCTGCGCAATATCCATTCCATCAACCCCCACACAA
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTGTTGTCACCTTGACCTCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTCCTGGCAGTTTCGATTTAAAGTGTAAACTTGCATATGCTGTGCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAG
TGATCCTTCGAAGCTTTCTCAAGGGGTTTGGCTATGTGCTGCCATCATTCACTCATTCACTC
CTTGCCTGGATTGAGACGTTGGCTGGATTCAAAGTGTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTCAGGATGCTCAGAGAGGGCAGCACTTACCTGGTGGCTTCTGATG
GTCAGTATTACCCCCCTCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAACAGGACAGT
GAGAACCACTTTAGAACTATGAGTACTACTTTGTTAAATGTGAAAACCTCACAGAAAGTC
ATCGAGGCAAAAGAGGCCAGGCACTGGAGTCTCCCTGTCGACAGTAAAGTGTGAATGGTAC
CACTGCTGGCTTATTGAACAGCTAATAAGATTATTGTAATACCTCACAAACGTTGAC
CATATCCATGCACATTAGTTGCTGCCGTGGCTGGTAAGGTAATGTCATGATTCACTCTCT
TCAGTGAGACTGAGCCTGATGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTAGTAAGCAAGATAACCTTTATTCAATTAC
AGAATGGAATTTTTGTTCATGTCAGATTATTGTATTCTTTAACACTCTACATT
TCCCTGTTTTAACTCATGCACATGTCCTTGTACAGTTAAAAGTGTAAATAAAACTG
ACATGTCATGTCAGTTATTCTGTTGCATTATGTCATGGCCTGAAGTGTGG
CTTGCAAAAGGGAGAAAGGAATTGCGAATACATGTAAGTCACTGAC
TTATCATGAAATCATGTTTCTCTGATTGTTCTGAAATGTTCTAAACTCTTATTGATATT
ACAAATGACTTAAACCATTCAATCATGTTCTTGCCTGAGCCAATTCAATTAAAGTGAAC
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLPOEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCCTGGGGCCGACCCGCCAGGAAAGACTGAGG
CCGC GG CCTGCCCGCCGGCTCCCTGCGCCGCCGCCCTCCCAGGACAGAAGATGTGCTCAG
GGTCCCCTGCTGCTGCCGCTGCTCTGCTACTGGCCCTGGGGCTGGGTGCAAGGGCTGCCAT
CCGGCTGCCAGTGAGCCAGCACAGACAGTCTTGACTGCCGCCAGGGACCGACGGTGCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTGAGAACGGCATCACCAGTGCAGCC
AGGCAGCTTGCGGCCCTGCCGGGCTGCAAGCTCTGGACCTGTACAGAACAGATGCCAGCC
TGCCCAGCGGGCTTCCAGCACTCGCCAACCTCAGAACCTGGACCTGACGGCAACAGGCTG
CATGAAATACCAATGAGACCTCCGTGGCTGCGGCCCTGAGCGCCTCTACCTGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTGACCCGCCCTCTGGAGCTCAAGCTG
AGGACAACGAGACTGCCGGCACTGCCCGCCTGCGCCTGCCGCCCTGCTGCTGCTGGACCTCAGC
CACAAACAGCCTCTGGCCCTGGAGGCCGATCTGGACACTGCCAACCTGGCAACTGGAGGGCTGCC
GGCTGGCTGGGGCTGCAAGCAGCTGGACGAGGGGCTCTCAGCGCTTGCACCTCCACGACC
TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCACCTGTGATCCAGGCTCCGGGGCTGACG
CGCCTGCCGGCTGGCCGGCAACACCCGATTGCCAGCTGCCGCCAGGACTGGCCGGCTGCC
TGCCCTGCAGGAGCTGGATGTGAGCAACCTAACGCTGCCAGGGCCCTGCCCTGGGACCTCTGGCC
TCTTCCCCCGCTGCCGCTGCTGGCAGCTGCCGCAACCCCTTAACACTGCGTGTGCCCCCTGAGC
TGGTTGGCCCTGGGTGCGCAGAGGCCACGTCAACTGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCCCCAAGAACGCTGGCCGGCTGCTCTGGAGCTGACTACGCCGACTTGGCTGCCAG
CCACCAACACCACAGCCACAGTGCCAACACAGGAGGCCCTGGTGGGAGGCCACAGCCTTGCT
TCTAGCTGGCTCTACCTGGCTAGCCCCACAGGCCGGCACTGAGGCCACGCCCTGCTCA
CACTGCCCAACCGACTGTAGGGCTGTCCCCCAGGCCACGTGCCACCGTCCACCTGCTCA
ATGGGGCACATGCCACCTGGGACACGGCACCTGGCTGCTGTGCCCCGAAGGCTTACG
GGCCTGTACTGTGAGAGCCAGATGGGAGGGACACGGCCAGGCCCTACACAGTCAGGCCAG
GCCACCAACGGCTCTGACCCCTGGGATCAGGCCGGTGAAGGCCACGCCCTGCCGTGGGCTG
AGCGCTACCTCCAGGGGAGCTCCGTGAGCTCAGGCCCTGGCTCACCTATCGCAACCTATCG
GGCCCTGATAAGGGCTGGTGAACGCCCTGGCTGAGCTGCCCTGCCCTGCCGTGAGTACACGGTCA
GCGAGGAGGCTGCCGGGAGGGCCATACACCCCTGCCCTGAGCTGCCCTGCCGTGCTGCC
CAGGGCCGCGAGGGCAACCTGCCCTGAGCTGCCCTGAGCTGCCCTGCCGTGCTGCC
GCTGGCTGCCGGTGGGGCAGCTACTGTGTCGGCGGGGGCGGGCATGGCAGCAGGGCTCAGG
ACAAAAGGGCAGGTGGGGCAGGGCTGGGCCCTGGAACTGGAGGGAGTGAAGGTCCCCCTGGAG
CCAGGCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGGGCTGAGTGTGAGGTGCC
ACTCATGGCTTCCCAGGGCTGGCTCCAGTCACCCCTCCACGCCAAAGCCCTACATTAAGCCA
GAGAGAGACAGGGCAGCTGGGCCGGCTCTAGCCAGTGAAGATGCCAGGCCCTCTGCTGCC
ACACCAACGTAAGTCTCAGTCCAACCTGGGATGTGTCAGACAGGGCTGTTGACCAAGCT
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGAGGCC
CTAACGTCCCCAGAACCGAGTGCTCATGAGGACAGTGTCCGCCCTGCCCTGCCAACGTGCA
CCTGGGCACGGCGGCCCTGCCATGTGCTGGTAACGCACTGCCCTGGCTCTGCTGGCTCTCCAC
TCCAGGCCACCCCTGGGGCCAGTGAAGGAAGCTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGACTCTAGCTTGCCCCAGGAAGCGAAGGAACAAAGAAACTGGAAAGGAAGATGC
TTTAGGAACATGTTGCTTTAAATATATATTTATAAGAGATCCTTCCATTATTCT
GGGAAGATGTTTCAAACTCAGAGACAAGGACTTGGTTGTAAGACAAACGATGATGAA
GGCTTTGTAAGAAAAAATAAGATGAAGTGTGAA

FIGURE 16

MCSRVPLLLPLLLLLALGPGVQGCPSCQCSPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGQLLQLSQQNQIASLPSGVFQPLANLSNLDTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILDtanve
ALRLLAGLGLQQLDEGLFSRLRNLDLVDSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLPGDLSGLFPRLLLAAARNPNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTGLIEPVSPSTSRLVGLQRYLQGSSVQLRSRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPLGPGRVPEGEAACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAAYCVRRGRAMAAAQDKGQVPGAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFPGPGLQSPFHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCAGGCGGCGGTGGCTGAGTCGTGGCAGAGCGAAGGCACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGTGCGGTGCTGAGCTGGCTCGCGTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTAGATTCAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGATTCAAAGAAATCTGA
ATTAGAACCTCTATTCAAGAACAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGACTCCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAGTA
CGGAAACCAGCTTGACGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGATGGCAGACTGTGGTGTG
CTACAACTATGACTACAAAGCAGATGAAAAGTGGGGCTTGTGAAACTGAAGAACAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGAATGAAAATCTTAATGAAAG
CAATAAGAAAAGCCAAAAAGAGAACATATCGGTATCTCAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATGCTCTTATTGGTATTACTGCCACAGAATATCCAG
GCAGCGAGAGAGATTTGAGAACAGCTGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTGG
CTTCTGTATGCCTCTGGACTGGTGTAAATTCAAGTCAGGCAAAGGCTTTGTATATTACAT
TTGGAGCTTGGGGCAATCTAATAGCCCACATGGTTGGTAAGTAGACTTTAGTGGAAAGGCT
AATAATATTAACATCAGAAGAATTGGTTATAGCGGCCACAACCTTTTCAGCTTCATGATC
CAGATTGCTGTATTAAGACCAAATATTCAAGTGAACCTCCTCAAATTCTGTAAATGGATAT
AACACATGGAATCTACATGAAATGAAAGTGGTGGAGTCACAATTCTTCTTAAATGGATAT
TTGGCTGATTGCCCTAAAAAGAGAGATCTGATAAATGGCTCTTTAAATTCTCTGAGTTG
GAATTGTCAGAATCATTTCACATTAGATTATCATAATTAAATTCTTCTTAAATTGCAACATGC
AAATTGTAATGGTGGCTATAGAAAAACACATGAAATATTACAAATTGCAACATGC
CTTAAGAATTGTTAAATTCAAGGAGTTATTGTCAGAATGACTCCAGAGAGCTACTTCTG
TTTTTACTTTCATGATTGGCTGTCTCCATTCTGGTCAATTGCTAGTGACACTGT
GCCTGCTCCAGTAGTCTCATTTCCATTGCTAATTGTTACTTTCTTGTCAATTGG
AAGATTAACTCATTAAATAAAATTATGCTAAGATTTTTTTTTTTTTTTTTTTTTTTTTTT
AAA

FIGURE 18

MRVRIGLTLLCAVLLSLASASSDEEGSQDES LDSKTTLSDESVKDHTAGR VVAGQI FLDSEEESEL
ESSIQEEEDSLKSQEGERVTEDISFLESPNPNENKDYEEP KVKRKPALT AIEGTAHGE PCHFPFLFLDK
EYDECTSDGREDGRILWCATTYDYKADEKWGF CTEEEEAKRRQM QEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYL PQNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTGFALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAAGTTAAGCCCATTCTGCAGTGGAAATTGATGAACATAGCAAGAGGACACCATTCTT
GTATTATAACAAGAAAGGAGTGACTTACATCACACACAGGGGGAAAAATGCTTTGGGTGCTAGG
CCTCCTAATCCTCTGGTTCTGTGGACTCGTAAGGAAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTCATCACTGGATGTGACTCGGGCTTGGAAACTTGGCAGCCAGAACCTTGT
AAAAAGGGATTTCATGTAATCGCTGCCGTGACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTGGGGAGAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGT
CCCGCGTGTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAAATGTCTCCAGTGTGGAGGTGCGCTTGCATCGTGGAGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTGGCAGATCCAGTAAAGGTAAATTGAAAAAAAC
TCGCCATTGGGAGCAGCTGTCTCCAGACATCAAACAATATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAAACTGAAAGGCAATAATCCTATGTGAACATGGACCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCAACAAGTCTTCCCTAAGACTCATTATGCCGTGGAAAGATGCCAAA
TTTCTGGATAACCTCTGTCTCACATGCCAGCAGCTTGCAAGACTTTTATTGGTAAACAGAAA
GCAGAGCTGGCTAATCCAAGGCAGTGTGACTCAGCTAACCAATGTCTCCTCCAGGCTATGA
AATTGGCCGATTCAGAACACATCTCCTTCAACCCATTCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGTCTATTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGTATCCAGGGT
CCCTGCTCAAGTTCTTGAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCT
GTATTTAGGCTTGGCTGCTGGTGTATGTAAGGGAAATTGAAAGACTTGGCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCATGCTTATGGCCCCAGCATTACAGTAACTTGTGAATGTTAAGT
ATCATCTTATCTAAATATTAAAGATAAGTCAACCCAAAAAA
AAAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLINNNAGVPGVLAPTDWLTLEDY
REPIEVNLFGGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEFYIEKSLDKLGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAAKGDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGC GGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGCG
CACTCGTTCCAGCACCTCAACACGGACTCGGACACCGAAGGTTCTTCTTGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATACAATTGA
CATTCA GAAATATATTCCATGCTATCAGTTTAGCTTATAATTCTCAGGCGAAGTAATG
AGCAAGCACTGAAGAAAATATTCAAATGTCAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTCA GATCAGATCATGACGTTAGAGAGAGGCTGCTTCACAAAAACTTGCA GGAGCATT
TTCAACCAAGACCTTGT TCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTACTC
ATCGACTGGAACATCCTTATATAAACCTCAAAAAGGACTTTTACAGGGTACCTTAGTGGTT
GCCAATCTGGCAGTGTCTGAACA ACTGGGTATAAAACTGTATCAGGTTCTGTATGTCAC TGG
TTTAGCCGAGCAGTACAAACACACAGCTCTAAATTGGAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCAAGAGGAATTAAAGAGTATATGCA AAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAACAGATTAAACAGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTAGGCAGCAAGAGAGAAGAACATCCAAAAGACCTCAGGAGA
ACATTTCTTGTCA GG CATTACGGACCTTTTCAAACTCTGAATTCTTCA TTGATGTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAGTAGCTGTA ACTACAACCACATCTGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATT CCTGAAGCTAGTCCAGCTAGTACACC
AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAACATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTGGTGAATATTACGGTCTC
CTACATTTTGATCTTAACCTTACAAGGAGATTTTTATTTGGCTGATGGTAAAGC CAAAC
ATTCTATTGTTTACTATGAGCTACTGAGTAAGTCATTGTTTACTATGTT CACC
TGTGAGCTAATACACAGATAACTCTTAGTGCATTACTCACAAGTACTTTCAACATCA
GATGCTTTATTCCAAACCTTTTACCTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATTCTTAGAATTGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
GAGACCATGTCTATTAAAAAATGGAAAAGCAAGAATAGCCTTATTTCAAAATATGGAAA
GAAATTATGAAAATTCTGAGTCATTAAAATTCTCCTTAAGTGATACTTTTAAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKNNVGWYKFRRHSDQIMTFRERLLHKNLQEHSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVGSCMSTGFSRAV
QTHSSKFFEDGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHLDVVDNLTL
MVEHTDipeASPASTPQIIKHKA LD LDDRWFQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGGAGCGGACCAGCGCAGGGCAGCCCAA
GCAGCGCCAGCGAACGCCGCCGCCACACCCCTCGCGTCCCAGCGCGCCTGCCACCCCTCCCTCTTCCCC
GCGTCCCCGCCCTCGCGGCCAGTCAGCTTGCCGGTTCGCTGCCCGCGAAACCCCGAGGTCAACAGCCCGGCCCT
GCTTCCCTGGGCCGCGGCCCTCACGCCCTCCTCTCCCGCTGCCCTCTCCACCAACTCCAACCTCTTCTCCC
CCGAGGGCCAGCTACTTTGCCCGTCTCTCCGCCCTGCCCTCTTCCACCAACTCCAACCTCTTCTCCC
TCCAGCTCCACTCGCTAGTCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGGCCGCTTCCCGTCCGGTCCAAA
GGTGGGAACCGCTCCGCCCGCCCGCACCAATGGCACGGTTCGGCTTGCCCGCCTCTCGCACCCTGGCAGTGCTC
AGGCCGCCGCTGCTGGCTGCCAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCAGCTTACGTGCTCAAAGGCTTC
AACAGAACGATGCCCTCCACGAGATCAACGGTATGCAATTGAGATCTGTCCCCAGGGTTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACGCCGCAAAAGTAAAGATGATTCAAAGTGTGGTCAGCGAACAGTGAATCATTTG
CAAGCTGTCTTGCTTCACTACAAGAAGTTGATGAATTCTCAAAGAACTACTGAAAATGCAGAGAAATCCCTG
AATGATATGTTGTAAGACATATGCCATTATACATGCAAAATTCTGAGCTATTAAAGATCTTCGTAGAGTTG
AAACGTTACTACGTGGGGAAATGTAACCTGGAAGAAATGCTAAATGACTTCTGGGCTGCCCTCTGGAGCGGATG
TTCCGCCCTGGTAACCTCCAGTACCAACTTACAGATGAGTATCTGGAATGTTGAGCAAGTATAACGGAGCAGCTGAAG
CCCTTCCGGAGATGTCCTCGCAAATTGAAAGCTCCAGTTACTCGTGTCTTGTAGCAGCCGTACTTCGTCAAGGC
TTAGCGGTGCGGGAGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGTCTCGTGAATGCAACTACTGCTCAAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGGATCTGATTTGAATGGAACATTCTAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTCAACATTGAATCGGTATGGATCCCCTCGATGTGAAGATTCTGATGCTATTATGAACATGCAGGAT
AAATAGTGTCAAGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGTGGACGAATTCTCGT
TCCATCTCTGAAAGTGCCTCAGTGTGCTCGTCAACATCACCCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCGACTGGTACTGATGTCAGGAGAAACTGAAACAGGCAAGAAATCTGGCCTCCCTCCAGCAAC
GTTTGCAACGATGAGAGGATGGCTCAGGAAACGCCATGAGGATGACTGTTGGAATGGAAAGGCAAAAGCAGGTAC
CTGTTGCACTGACAGGAAATTGATTAGCCAACCAGGGCAACACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCTTCGTCAAATCATGGCTTCTCGAGTGTGATGACAGCAAGATGAAGATGCTACAAATGGAACGACGTG
GACTTCTTGATATCAGTGTGAAAGTAGTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGAGAGTGCCAATGAGAAAAGCCGACAGTGTGCTGGTGTCCCTGGGCA
CAGGCCACCTCCCTCACTGCTCTGCATCTGGTTCTGGTATGCAGAGAGAGTGGAGATAATTCTCAAACCTGAG
AAAAAGTGTTCATCAAAAGTTAAAGGCACCGTTATCACTTTCTACCATCCTAGTGACTTGTCTTTAAATGAA
TGGACAAACATGTACAGTTTACTATGTGGCACTGGTTAAGAAGTGTGACTTGTCTCATTCACTGTTGG
AGGAAAAGGGACTGTGCATTGAGTTGGTCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTGTGATTCTACTCTATTATTTGTTGATGTTTTCTCATTTGTTGTGGTT
TTTTTTCCAACTGTGATCTCGCCTTGTCTTACAAGCAAACCAAGGGTCCCTTGGCACGTAACATGTACGTATT
TCTGAAATATTAATAGCTGTACAGAACAGGTTTATTATCATGTTATCTTATAAAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGVNVLEEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
QLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVSKVSVNPTAQCTHALLKMIYCSHCRL
VTVKPCNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFGCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSLPSNVNCNDERMAAGNGNEDDCWNGKGKSRYLFATVTGNGLANQNNPEVQDTS
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCATTATATTCTTCAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTCTAGGAGATGGCTCCAGGAAGGGGCCAAGAATGTGAGTGCAAAGATTGGTCCTGAG
AGCCCCGAGAAGAAAATTCA TGACAGTGTCTGGGCTGCCAAAAGAACAGCAGTGCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAAATGTCAGCTAAGAAGCTTGCTCTGCCTTGTAGGAGCTCTG
AGCGCCCACTCTCCAATTAAACATTCTCAGCCAAGAACAGT GAGCACACCTACCAAGACACTC
TTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTGTTGCTCTCTAGTGTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTAACCGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCC
AGCTAGTGTCAATTAAACCTAAATGCAATCAGGAAAGTAGCAAACAGAAAGTCAATAAATATTTT
AAATGTCAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLPLMLMSMVSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKKQCPKCDHFKGKVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACCTCCTGCAGCTGCTGGTGTGCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGCTGCTGGCAGCCCCGTGCAAAAGCTACTTCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGAAGATGGAGAGCAAGAAACGGGAGCTTCAGCCAG
ATAAAGGGCTTACAGGAGCCTCGGGAAAGTGGCCTACTGGAGCTGGCTGCCAACCGGAGC
CAACTTCAGTTCTACCCACCAGGCTCACCTGCCTAGACCCAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGCTGCACTCTGGTGTG
CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGGAGACTGAGACCCGGAGGTGTGC
TCTTTCTGGGAGCATGTGGCAGAACCATATGAAAGCTGGGCTTCATGTGGCAGCAAGTTTC
GAGCCCACCTGAAACACATTGGGATGGCTGCTGCCCTACCAGAGAGACCTGGAAGGATCTGA
GAACGCCAGTTCTCGAAATCCAATGGAACGACAGCCCCCTCCCTGAAGTGGTACCTGTTG
GGCCCCACATCATGGAAAGGCTGTCAAACAATCTTCCAAGCTCCAAGGCACCTGGCTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCCTATCTATCTCCACTGAGAGGGACCTA
GCAGAACATGAGAGAACATTGATGTACCACTACTAGTCCCTCTCCCCAACCTCTGCCAGGGC
AATCTCAACTCAATCCGCCCTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGATCCTCAACTGCAAGTTCTGGACTAGTCTCCAACGTTGCCTC
CCAATGTTGCCCTTCCCTCGTCCCATGGTAAAGCTCCTCTCGCTTCCCTGAGGCTACAC
CCATGCGTCTAGGAACGGTACAAAAGTCATGGTGCCTGCATCCCTGCCAACCCCCCTGAC
CCTCTCTCCCCACTACCACCTCTTCCCTGAGCTGGGGCACCAGGGAGAACAGAGATGCTGGG
ATGCCAGAGCAAGACTCAAAGAGGGAGGGTTTGTCTCAAATATTTTTAATAATAGACGAA
ACCACG

FIGURE 28

MDILVPILLQLLVLLTLPLHLMALLGCWQPLCKSYFPYILMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQS PRKVLQEVRVRLPGGVLF FWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTGGAGTGGGACTGGTCGGTCCCAGAAAGTCTCTCTG
CCACTGACGCCCATCAGGGATTGGGCCTCTTCCCCCTCCTCTGTGTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGGAGAAAGTGGGGATGGTA
AGAAAGCTGGAGATAGGAAACAGAACAGAGGGTAGTGGGTGGCTAGGGGGCTGCCTTATTAAA
GTGGTTTTATGATTCTTACTAATTACAAAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTTCCCTGTGTTCAATGTTGTAAAGATTGTTCTGTGTAAATATGTCTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAA

FIGURE 30

MLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCTTCAACTATACCCACAGTCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCCTATTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTGATAAAATGATGTTGACACCCCTCACCGAATTCTAAGTGGAAAT**CATGTCGG**
GAAGAGATAACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTGTCTTGGCCATGATGTTACC
TTCAGATTCATCACCAACCCCTCGGTTCACATTTCATTGGTTATTTGGGATTGTTGTT
TGTCTGCGGTGTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGCCTGCTGGGTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGTCTCTGGTGGCTGTGCTGAGCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCAGCCAAGTGGAAATATAAGCCCCTTCGGCATTGGTACATGTGGCGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTGCGTGCCAGCAAATGACTATAGCTGGGCAG
TGGTTACTTGTATTCAACAGAAGTAAAAATGATCCTCTGATCATCCCATTCTGTCTCTC
TCCATTCTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTATTTTAATCTGTGGTAG
GATTCCGAGAACATTGTATGTACATGCAAACGCACTGAAAGAACAGCAGCATGGTCATTGT
CCAGGTACCTGTTCCGATGCTGACTGCTGTTCTGGTGTCTGACAAAATACCTGCTCCATCTC
AACCGAAATGCATATACTACAACACTGCTATTAATGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAAATCTTGCCAAGAAACTCAAGTCACCTTACATCTATTAACCTGCTTGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGTGGTGTGTTCACTGTTGGAGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGTGGCAGTCCCTCTGTTATTGGTAGCTTTGGCTACTTAGTAGC
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTGCTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAGCCCTACTTATGGATCAAGAAATTCTGAGTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGA**TAGATACCCATTAGGTATCTGTACCTGGAAAACATT**
TCCTTCTAAGAGCCATTACAGAAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATT
TTAAAAGACCTAATAACCCATTCTCCTCAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISLVLGILFVCGLWLYDYTNDSLIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDDPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMMQNALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTAINGTDFCTSAKDAFKILSKNSHFTSINCFGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNAQDQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTGGACCTCCCTGTTCTCCCTAGA
ATAATTGTATGGGATTGTGATGCAGGAAGCCTAACAGGAAAGAATATTCAATTCTGTGTTGGT
GAAAATTTTGAAAAAAATGCCCTCTCAAACAAAGGGTGTCAATTCTGATATTTGAGGAC
TGTTGTCTCACTATGAAGGCATCTGTTATTGAAATGTCCTGTTGCTGGTACTGGAGTAC
ATTCAAACAAGAACGGCAAGAAGATTAAAGGCCAACGTTCACTGTGCCTCAGATCAACTGC
GATGTCAAAGCCGAAAGATCATCGATCCTGAGTTCAATTGAAATGTCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGATGCATCCTACTCCAGTGTGTGGCGCTGCCG
TACACAGTGGTGTGTTGATAATTCAAGGAGGGAAATACTTGTCGGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCAACGTTATCCCTACCACGATGGAGAGAAATCCTT
TATCGTCTTAGAAAGTAAACCAAAAAGGGTGTAAACCTACCCATCAGCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCAACGAGGTGAGACCACAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACATGCACAGCCGGTCACTCTGATGCAGCTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
CACCTTGCCAAGGCCATCCCCTCTGCTGTTCTACCACCAGATCCCCAGACCAACATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCACTGCCACCTACACAAGCAGCCAAAAGGCC
AGAGCTGATCCAGGTATCCAAGGCAAGATCCTCAGGAGCTGCCCTCCAGAAACCTGTTGGAGC
GGATGTCAGCCTGGACTGTTCCAAAAGAAGAATTGAGCACACAGTCTTGGAGCCAGTATCCC
TGGGAGATCCAACCTGCAAAATGACTTGTGTTTATTGATGGAGCACCAGCATTGGCAAA
CGGGCATTCCAATCAGAGCAGCTCTGGCTGATGTGCCAAGCTTACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCAGTATGGAGACAAACCTGCTACTCACTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTCTAAT
GTAGGTCGGGCCATCCTTTGTGACCAAGAACCTCTTTCCAAGCAATGGAACAGAACAGCGG
GGCTCCAATGTGGTGGTGTGATGGTGGATGGCTGCCACGGACAAAGTGGAGGAGGCTCAA
GACTTGCAGAGAGTCAGGAATCAACATTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGCCCAACTTGCACAAACAAAGGCCGTGTCAGAACACAAACGGCTCTACTC
GCTCCACGTGCAAGAGCTGGTTGGCCTCCACAAGACCCCTGCAAGCCTCTGGTGAAGCGGGTCTGCG
ACACTGACCGCCTGGCCTGCAAGCAAGACCTGCTGAACATTGGCTCATCGAC
GGCTCCACGAGTGTGGGGACGGCAACTTCCGACCGCTCCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTCGACACGGACACGCGCATCGGGCCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTGGGTTGCAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTGGTGGCACAGCAGCGACGGGCTGCCATCAACTTCCGCTGGAGCAGCTCTCAAGAA
GTCCAACGCCAACAGAGGAAGTTAATGATCCTCATCACCGACGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGTACACCTATGCGATAGGCGTTGCCCTGGCT
GCCCAAGAGGAGCTAGAAGTCAATTGCCACTCACCCGCCAGAGACCACTCCCTTTGTGGACGA
GTTTGACAACCTCCATCGATATGTCCTCAGGATCATCCAGAACATTGTAAGAGTTCAACTCAC
AGCCTCGGAACTGAATTCAAGAGCAGGACAGCACAGAACGTTCTACTAAGTACGTGTT
GGACCAACCCACCGCTTAATGGGCACGCACGGTGCACTAACGTTGGCAGGGCATGGAGAAC
AAATGTTGTTATTATTCTTGCATCATGCTTTCATATTCCAAAACATTGGAGTTACAAAGA
TGATCACAAACGTATAGAATGAGCAAAAGGCTACATCATGTTGAGGGTGTGAGGATTTACAT
TTTGACAATTGTTCAAAATAAATGTTCCAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
AGCTTTGTGAGATTTAAAGTGTATTCTGATTTGAACTCTGTAACCCCTCAGCAAGTTCT
TTTGTCATGACAATGTAGGAATTGCTGAATTAAATGTTAGAAGGATGAAAAAATAAAAAAAAA
AAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINC DVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGYSNGVQSLSLPRWR
ESFIVLESKPKKGVTVPSALTYSSSKPAAQAGETTKAYQRPIP GTTAQPVTLMQLLAVTVAVA
TPPTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCIDLSFLIDGSTSIGKRRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTTLQPLVKRVCDTDRCLACSKTCLNSADIGFVIDGSSSGTGNFR TVLQFVTN
LTKEFEISDTDRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKR VGYWSGGTSTGAIA INF AQL
FKKS KPNKRKL MILITDGRSYDDV RI PAMA AHLKG VITYAIGVAWAQEELEVIATHPARDHSFF
VDEF DNLHQYVPRIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCCTTAGGAGGTGGCTGCCTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTTCTTCTGTTTCAGAGTAGTCACAACAGATCTGAGTGTTAATAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTCTATCTGGATTCCACAGTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTCCCGCTGGCTGCTATCACGTGGTGTCTCCGACTACTCACCCCGAGTGT
AAGAACCTTCGGCTCGCGTCTGAGCTGCTGGAGGCCTCGCTCTGGACTGTCCCTGGAGTA
GGATGTCACTGAGATCCCTAAATGGAGCCTCGCTGCTGACTCCTGAGTTCTTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACGCGTACACTGGATGTACTTCTATGAGTATGAGCCGATT
CAGACAAGACTTCACTTCACACTCGAGAGCATTCAAACTGCTCTCATCAAATCCATTCTGGTCATT
TGGTGACCTCCCACCCCTCAGATGTGAAAGGCCAGGCCATTAGAGTTACTTGGGGTGA
TGGTGGGATATGAGGTCTTACATTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGA
ACCTGACCTGAAAACCATTATGGCATTCAAGGTGGTAACTGAGTTTGCCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTCATCAAATACTGGCAATTAGTGAAGTATCTTAAACCTAAACCACTCAGA
GAAGTTTCACAGGTTACCTCTAATTGATAATTATTCCCTATAGAGGATTTACCA
CTTACCCAGGAGTATCCTTCAAGGTGTTCCCTCAACTGCAGTGGGGTGGGTATATAATGTCCAGAGAT
TTGGTGCCAGGATCTATGAAATGATGGTCACGTAAAACCCATCAAGTTGAAGATGTTATGTCGGAT
CTGTTGAATTATTAAAAGTGAACATTCAATTCCAGAAGACACAATCTTCTTCTATAGAATCC
ATTGGATGTCGTAACTGAGACGTGATTGCAGCCCAGGTTCTCCAGGAGATCATCACTTT
TGGCAGGTCATGCTAAGGAACACCACATGCCATTAACTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTGTGAAAGTGTAAATAAGTAGGACTGTGGAAAATTCAAGGAGGTCACTGTGCTGGCT
ACACTGAACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGATTTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAAATAAGGAATTGGAGGTTTGCTAAAGAAATTAAATAGG
ACCAAACAATGGACATGTCATTGTAGACTAGAATTCTTAAAGGTGTTACTGAGTTATAAGCTCA
CTAGGCTGAAAAACAAACTGTAGAGTTATTGAACAATGTAGTCATTGAAGGTTTGCTA
TATCTTATGTGGATTACCAATTAAAATATATGTAGTTCTGTGTC
CTGAACAAAATTTCACCTGTTGGTCATTATAAAAGTACTCAACATGTTGCAGTATTCA
ATTATTTAAATTACTCACTTGTGTTTAAATGTTGACGATTCAATACAAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACTTCCAGTTACTTAACTGATCAGTTATTATTGATA
TTAATGTAAGTCATAGGTCAATTGCATATCAGTAATCTTGGACTTGTAAATTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAATCTGAAAAA

FIGURE 36

MASALWTVLPSRMSLRSILKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNCSSHQNPFVLVILVTSHPSDVKARQAIRVTWGEKKSSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIHQDFLDTYNNLTLKТИMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNNSYRGFYQKTHISYQEYPFKVFPPYCSDLGYIMSRLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCGCCGGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTCTCAGCTCCTTCATC
TCGTCCCTGCCAAGAGAGTACACAGTCATTAAAGAACCTGCCCTGGAGCAGAGTGAATATCATGTCGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTGCCCCGGAAAGAGGGAAAGCTGTGGGTTATACCATCCCTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCCTGATCCACCCAGGTTGACTACATCTTGAAACATGCAAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTGGATGACTCTATGTGAAGGGGTCTACTGTGCAGAGTGCCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTGTTGAAAGCTATCCCCTAAATGCTCACTGTGAATGGACCATTCATGC
TAAACCTGGGTTGTCATCCAATTAAGATTGTCATGTTGAGTCTGGAGTTGACTACATGTCCTGAGTATGACTATGTTGAG
GTTGAGTGGAGACAACCGCGATGCCAGATCATCAAGCGTGTCTGGCAACGAGCGGCCAGCTCTATCCAGAGCATAG
GATCCTCACTCCACGTCCTTCCACTCCGATGCCAAGAATTTCAGCGTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCCTCATCCCTTGTCTTCCATGACGGCACGTGCGCTTGCACAGGCTGGATCTACAAAGTGTGCCCTGCTTGGCAGGC
TATACTGGGCAGCCTGTAAGAAACTCTCTTGAAGAAAAGAAACTGCTCAGACCCCTGGGGGCCAGTCATGGTACAGAAAA
TAACAGGGGGCCCTGGGCTTATCACGGACGCCATGCTAAATTGGCACCGTGTCTTCTTGTAAACAACCTCTATGT
TCTTAGTGGCAATGAGAAAAGAACATTGCCAGCAGAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTTCAGACCTGGTGAGAAGGAGAGTTCTCCGATGCGAGTTCACTGCAAGGAGACACCATTACACCAGCTAT
ACTCAGCGGCCTTCAGCAAGCAGAAACTGCAGAGTGCCCTACCAAGAACGGCAGCCCTCCCTGGAGATCTGCCATGG
ATACCAACATCTCATACCCAGCTCCACTATGAGTCATCTACCCCTCTACCGCCCTGGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCATCCTGCATCCCTATGCGGGAAATTGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCGTGGCAGGCCATCTACAGGAGGACAGCGGGGTGCACTGACGGCAGCCCTACAAAGGGAGC
GTGGTTCTCTAGTGCAGCGGTGCCCTGGTGAATGAGCGACTGTGGTGTGGCTGCCACTGTGTACTGACCTGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTGGGAAATTCTACCGGATGATGACCGGGATGAGAACACCAC
AGAGCCTACAGATTCTGCTATCATTCTGCATCCAACTATGACCCCATCCTGCTGATGCTGACATGCCATCCTGAAGCT
CCTAGACAAGGCCGTATCAGCACCCAGTCCAGCCATCTGCCTCGCTGCCAGTCGGGATCTCAGCACTCCCTCAGGAG
TCCCACACACTGTGGCTGGCTGGAATGTCCTGGCAGACGTGAGGAGCACCATGGCATTCCAGTGTGACTGATAACATGTT
TGGTCAGTGTGGGACTCGCTGTGAGGAGCACCATGGCATTCCAGTGTGACTGATAACATGTT
CTGTGCCAGCTGGGACCCACTGCCCTCTGATATCTGCACTGCAGAGACAGGAGGACATGCCGCTGTGCTTCCGGGA
CGAGCATTCTGAGGCCAGCTGGCATTGATGGACTGGTCAGCTGGAGCTATGATAAAACATGCAAGCCACAGGCTCTCCA
CTGCCCTCACCAAGGTGCTGCCCTTAAAGACTGGATTGAAAGAAATGAAATTGACCATGCTCATGCACACTCTTGAGAAG
TGTTCCTGTATATCGCTGTACGTGTCAATTGCGTGAAGCAGTGTGGCCTGAAGTGTGATTTGGCTGTGAACCTGGCT
GTGCCAGGGCTCTGACTCAGGGACAAAACACTGAGAAGGGTGAGTAGACCTCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAGATGCCAGGGCTTGCAGAGAAGTAACCTTCTCAAAGAACAGACCATATAACAAACCTCTCCA
CTCCACATGACCTGGCTTCCCAACTTCACTGATGACATGCCATCAGCTTGACCCAGGGAAAGATCTGGCTCATGAG
GCCCTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCCTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCT
TGTGTACAGGCCACAGTACAGTCTGGCTTCTCCCTCCCATCTCTGTACACATTAAATAAGGGTTGGCTTCT
GAAGTACAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAEWNIMCRECCYEYDQIECVCPEGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRGENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSNE
KRTCQQNQEWSGKQPICIKACREP KISDLVRRVLPQMVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPGMYQHLHTQLOQYECISPFYRRLGSSRTCLRTGKWSGRAPSCIPCGKIENITAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDREKTIQSLQISAIILHPE NYDPILL DADIAILKLLDKARI STRVQPI CLAASR
DLSTS FQESHITVAGWNVLADV RSPGFKN DTLRSGVVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPS DICTAETGGIAAVSF PGRASPEPRWHLMGLVWSYDKTCSHRLSTAFTKVL PFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTCTCTACATCCTCTCATCTGAGAATCAGAGAGCATAATTCTTCTACGGGCCGTGATTATTAAACGTGGCTTAATCTGAAGGTTCTAGTCAAATTCTTGTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTAAAGGAGCTTGGCTGGTTGGCCCTTGAGCTGACAGAAGGTGGCCAGGGAGATGCAGCACACTGCTCGAGAATGAAGGCGCTTCTGTG
TGGTCTTGGCTTGGCTCAGTCCTGTA~~CT~~ACTACATTGACAATGTGGCAACCTGCACTTCTGTATT~~CAGA~~ACTCTGTA
AAGGTGCCCCACTACGGCCTGACCAAAGATAGGAAGAGGCCTCACAGATGGCTGCCAGACGGCTGTCGAGCAGCC
TCACAGCCACGGCTCCCTCCCAGAGGTTCTGCAGTGCCACCATCTCTTAATGACAGACGAGCCTGGCTAGACAA
ACCTGCTTAGTGTCTCGCAGAGGACGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGACTA
GGGCACGGGCTTTGAGAGATCCACTATTAGCAGATCATTAAAAAAATTCAGGCTTGTAGTGTCTTCGAA
GGACAAAGAGCGGGAGTGCAGTGCACCAGCCAGGGAAATTCTGAAAACACCAGTGCCTGGAG
TCTTCTCAAGGTTGTA~~CC~~ACTGATTCCAGATGGTGAATTACAGCATCAAGATCAATCGAGTAGATCCAGTGA
GCCTCTCTATTAGGCTGGGGAGGTAGC~~GA~~ACCCCCACTGGT~~CA~~TATCATTATCCAA~~CA~~CATTATCGTATGGG
TGATGCCAGAGACGCCGGCTACTGCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCTC
ACAACTACGCTGTGCGTCTCC~~G~~CCAGGTGCTG~~GG~~CTGACTGTGATGCGTGAAGAGATTCCCGA
GCAGGAACAATGGACAGGCCCGATGCC~~T~~ACAGACCCAGATGACAGC~~TT~~ATGTGATTCTAACAAAAGTAGCC
CCGAGGAGCAGTGGAA~~AA~~ACTGGTGC~~GA~~CAAGGGATGAGCTGGTTT~~CA~~TCTCAATGTGCTGGATGGC
GTGTGGCATATGACATGGT~~CAG~~TTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCA
GCCAGAAAGTGC~~GG~~CTCATCTGATT~~CAG~~CCAGTGAAGAGACGTCTCACCTCGTGT~~CC~~CCAGGT~~GG~~CAGC
GGAGCCCTGACATCTTCAGGAAGCCGGCTGGAACAC~~CA~~ATGGCAGCTGGT~~CC~~CCAGGGCAGGGAGAGGAGCAACA
CTCCCAAGCCCTCCATCTACAATTACTGT~~CAT~~GAGAAGGTG~~TA~~AT~~TC~~AAAAGACCCGGTGAATCTCTG
GCATGACCGT~~CG~~CAGGGGAGCATCACATAGAGAATGGATT~~TG~~CTATCTATGT~~CAT~~CAGTGTGAGCCGGAGGAG
TCATAAGCAGAGATGGAGAATAAAACAGGTGACATT~~TG~~GAATGTGGATGGGTCGA~~ACT~~GACAGAGGT~~CAG~~
GGAGTGAGG~~CAGT~~GGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTGGAGTCAAGACTATGAGC
CCCAGGAAGACTGCAGCAGCC~~CC~~AGCAGCC~~TG~~GACTCCAACCACACATGGCC~~CC~~ACCCAGT~~GACT~~GGT~~CC~~CCATCT
GGGT~~CAT~~GTGGCTGGAATTAC~~CC~~GGT~~GCT~~TATAACTGTAAGAGATATTGTATTACGAAGAAACACAG~~TG~~GAAGTC
TGGGCTTCTGCATTGTAGGAGGTATGAAGAATACAA~~TG~~AAACAAACCTTT~~T~~CATCAA~~AT~~CCATTGTTGAAGGAA
CACCAGCATACATGATGAGAATTAGATGTGGT~~GAT~~ATTCTCTGCTGTCAATGGTAGA~~ACT~~ACATCAGGAATGA
TACATGCTTGGCAAGACTGCTGAAAGAAC~~TTAA~~AGGAAGAATTACTCTA~~ACT~~ATTGTTCTGGC~~TGG~~CACTT
TTTATAGAATCAATGATGGGT~~CAG~~AGGAAAACAGAAAAATC~~AA~~AATAGGCTAAGAAG~~GT~~GAAACACTATATTATC
TTGTCAGTTTATATTAAAGAAGAACATATTG~~TA~~AAAATGT~~CAG~~GGAAAAGTATGATCATCTAATGAAAGCCAGT
ACACCTCAGAAAATATGATT~~CC~~AAAATTA~~AA~~ACTACTAGTTTT~~T~~CAGTGTGGAGGATTCTCATTA~~CT~~ACTCTAC
AACATTGTTATATT~~T~~CTATTCAATAAAAGCC~~TAA~~ACAA~~CT~~AAATGATTGATTGTATACCCACTGAATT
CAAGCTGATTAA~~TT~~AAAAT~~TG~~GTATGCTGAAGTCTGCCAAGGGTACATTATGCCATT~~TT~~AA~~TT~~ACAGCT
AAAATTTTAA~~AT~~GCA~~TG~~CTGAGAAGC~~TT~~GCTCATCAA~~AC~~AGAATAAAT~~TT~~TCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDCASLTATAPS
PEVSAATISIIMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIQHIIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNNGQAPDAYRPRDSFHVLNKSSPEEQLGIKLVRKVDEPGVFIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPEAAHLIQASERRVHLVSRQRSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDRGR
IKTDILLNDGVELTEVSREAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCDDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTCAGTTGTCATCAAGTCGAATCAGATTGGAAAAGCTCAACTTGAAGCTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATAATTACAGTAATAAAACATGGGCTTCAACCTGACT
TTCCACCTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTGTGCCTGACAGTGGTGGTGGC
CACCACTAAGTAACTTCTGGTGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCCCTCATTTGGGGAAAGGGAAAAACTCTGACTAATGAAGCATTCCACGAAGAAGGTAGAACCT
GACAACGTCCCTCTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTCAAACCAAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGAAAG
CTTTACAGAGGGTCGCCATCCTCGTCCCCACCGGAACAGAGAGAAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCTCCACCAGGCTGAAGG
TAAAAAGTTAATCGAGCAAACCTCTGAATGTGGGCTATCTAGAACCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGGAACAGCACTGGTACAGGTTACGTTACAGTGGATATTG
GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTCAAGGTGAATGGATTCTCTAACAACTACTGGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCAAAGAATGAAAATTCCCGCCCCCTG
CCTGAAGTGGTAAATATAACATGGTCTTCCACACTAGAGACAAAGGAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACGAGTCAGGATGGAGAACAGATGGTTGAGTAGTTGTTCTTATAAAAT
TAGTATCTGTGGAACACAATCCTTATATATCAACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTGTGTTGGAAAGAACTGATTCTTGTGCAATAATTGGCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTCCTTTGTATTTCT
TAGCAGAGCTCTGGTGTAGAGTATAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTGAAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAAGGAGATTATTAAATTGAAAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGCTGAGAGAACAGAGTTGTTCTGCTCCAAGGTAGAA
AGGTACGAAGATAACAATCTGTTATTCAATTCTGTACAATCATCTGTAAGTGGTGGTGTAGGT
GAGAAGGCGTCCACAAAGAGGGAGAAAAGGCAGCAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCCTCAGTGATGCCACCAGAGAATACATTCTTATTAGT
TTTAAAGAGTTTGTAAAATGATTTGTACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEGKKFNRAKLLNVGYLEALKEENWDCFIFHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRSGYFGVTALSREQFFKVNGFSNNYWGWGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGACAGCCAGACAGACGGCACGATGGACTGAGCTCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTCCCACAACAG
ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCAGAAGGAGGCAGACACCCACTTCCCCATCTGCATTCTGCTGCGGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCCCCTCCC
TTCCCTATTTATTCTGCTGCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLASLTGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATTGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCA~~G~~CAGCCTCTGGACCGTGAAAGAGCTGGTCGGT
TCCGGTGGTGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTATTGTCG
GACCTTCAACACAACCCCTCTGTACCACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAA~~A~~CTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGATATA~~C~~AGCTCATCACTCCAGCAGCCCTCCACCC
GGAGTACGTGCTGCATGTCTACGAGCACCTGCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCATAATGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGT~~G~~ATATGACCTTCATCTGC~~G~~TTGCCAGGAACCTGTCAGCAGAA~~A~~CT
TCTCAAGCCCCATCCTGCCAGGAAGCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCCTGTTGGT~~G~~CCCTCCTGCTCAGTCTTTGACTGGGCTATTCTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGTGGACATTGTCGG
AAACTCTAACATATGCC~~CC~~ATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAA~~T~~CCTAAAGGAAGATCAGCAAATACGGTTACTCCACTGTGGAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCTAAGTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLSKVKQVDSIVWTFTNTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIG
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLVPLLSSLFVLGLFLWFLKRERQEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTCTACTGCTGTTAGGAGTAGTTCAATGCGATACCTCTAATGTCAGCT
TAGTTGAGGAAGACCAATTTCTCAAACCCCATCTCTGTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAAATGCCTTGACAGCAAGAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTCTTCATCATTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTTAAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTAATTTCAATTGAAAAACATCAGTGACATTCATCCAGAATCCTCAA
CTTGCAGTGGTTTCAATGACTCTTGCACCTCCTACTGGTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAAACATAGG
CTTATCCACTTCTCAGTATTTTAGGTCTATTGCTTGGATTCTGGAGGTCTGTTGGGCT
CAGTCAGATAGTCATCGGTTCCTGGCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGGAAAAAATGTAAGTATCAGTAGTTGAAAAAAA
A

FIGURE 48

MTCCEGWTSCNGFSLLVLLLGVVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGQLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGACCCTGGCCCTTCCTACTGGCAGGCCTGACTGCCTTGGAGCCAATGACC
CATTTGCCAATAAGACGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGC GGAGGGCTCCTGGCATTGCTGGGATCGCGGAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACAGCACAGTCCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALEANDP FANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGCACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGCAGAGATGAAGTTCCAG
GGGCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGCAGTGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAGCACTGGGACAATATTGGGAGGCCCTGGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAGAGAACGAGCTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGCGC
AGCAGATGCTTGGCAACAGGGTCGGGAAGCAGCCCAGCTGGAAACACTGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCGCCGGCTCTGGCAGGG
GTGCCTGGCCACAGTGGTCTGGAAACTCTGGAGGCCATGGCATCTTGGCTCTCAAGGTGG
CCTGGAGGCCAGGGCAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGGATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGAATGGA
GGGCCACCAAACCTGGGACCAACACTCAGGGAGCTGGCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTCACGAATCCCCCACCCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGTGGCAGCAATGGTACAACAAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAG
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGCAGTGAGT
CCTCCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGAGCGGGAGGAAATGG
CATAAACCCGGGTGTGAAAAGCCAGGGATGAAGCCGGAGCGGGGAATCTGGATTAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTTGG
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCGAGCTGGGAGTGGAGGGTGAACGCTGTT
GGTGGAGTCATACTGTGAACACTCTGAGACGGCTCCTGGATGTTAACCTTGACACTTCTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGGATGCCATAAACAAAGGACCAGAGAACGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGAGCCCCCACACTCCCTTAA
AACACCAACCTCTCATCACTAATCTCAGGCCCTGCCCTGAAATAACCTTAGCTGCCCAACAA
AA
AA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVP
GHSGAWETSGGHGIFGSQGGLGGQQGQNPGLGTPVHGPNGNSAGSFGMNPQGAPWGQGGNGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGSGSQSGSSGSGSNGDNNNGSSGGS
SSGSSSSGSSGGSSGGSSGNSGGSRGDSGSESSWGSSTGSSSGNHGGSGGGNKHKGCEKPGNE
ARGSGESGIQGFRGQGVSNMREISKEGNRLGGSGDNYRGQGSSWGSQGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGGACAAGCTGCTCCGACAGAAGGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGACCTATGCCTCTATAACAACTGCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAACTGGTTTGGGTCACCTGGGCTGATCACTCCTACAGAGGAGGC
TTGAAGGACTGACCCAGATGTCGGCCACCTATTCCAGGGCTTACGGTATGGCTGGTCCCAT
CATCCCCCTCATCGTTTATGCCACCCGTACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCAAGGATAATCTCTTCATCAGGTCTGAAGCCCTGGCTGGAGAAGGGAACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCGCCATTCAACATCCT
GAAGTCCTATATAACGATCTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTGAGCACATCAGCCTCATGACCTGGACAGTCTA
CAGAAATGCATCTTCAGCTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTGTAGAGAAAAGAACGCAGCATATCCTCCAGCACATGGACTTCTGT
ATTACCTCTCCCAGTGACGGCGCGCTTCCACAGGGCTGCCCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGAGCGCGTGCACCCCTCCCACTCAGGGTATTGATGATTTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGTCTGTACAACCTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTCTGAAGGACCGCGATCTAAAGAGATTGAATGGGACGACCTGGCC
AGCTGCCCTCCTGACCATGTGCGTAAGGAGAGCCTGAGGTACATCCCCAGCTCCCTCATC
TCCCAGTGCTGCACCCAGGACATTGTTCTCCAGATGGCCAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGTCCATCACAACCCAACTGTGTTGGCCGATCTGAGGTCTACGACC
CCTTCCGCTTGTACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTATTCCCTTCTCCGCA
GGGCCAGGAACTGCATGGGCAGGCAGTCGCCATGGCGAGATGAAAGTGGCCTGGCGTTGAT
GCTGCTGCACTTCCGGTCTGCCAGACCACACTGAGCCCCGCAAGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGGGGTGGAGCCCTGAATGTAGGCTTGCAGACTTCTGAC
CCATCCACCTGTTTTGCAGATTGTCATGAATAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWLGRLPVAMSPWLLLLVVGSWILLARILAWTYAFYNNCRRLOC FPQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSDGRRFHRC
RLVHDFTDAVIRERRTLPTQGIDDFFKDKAKSKTLDFIGVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCQEVAELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPVEVYDPFRDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGTGTTGCACCTACCCGTGTTCTGCCT
TTTGGTGGATAACAAGGGACTTGCACTTATCTCTGCATTTGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCTGC
ATAATTCATGCCAGTTTATGAAGCTTGAAGGCACATGGACAGAACAGCTGGGACAGTTT
GTAACTATCTCGAAACCTCTGTCTTACAGACATGTGCCTTATCTTGCAGCAATGTGTTGCTT
GTGATTGAAACATTGAGGGTTACTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGAATCTCCTCATGTACCTGTTCTC
TCTGGATGTTGTCCCACTGAATTCCCAGAATACAAACCTATTCAACAGCAACAGCAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATGGCTCGAGGGGAGTGGAGCACCCAGCAGGCCAACATGCCTGTCTGCCCTG
TACGTGCCGTCATGGGAAGCCCACCCGAGTCAGACTTCAGACTTTGAGTCGAAGGGCTCCCTGCCAGCTGAAGTCC
ATTTCAAGCTCAGTGTCTCATCCCCCCCAGGAATTCTCACCTACCGCAGTGAAGCAGAAAATTGACAAGCT
GGAGATAAGGACCTTGATGGCAGCTAGACTTGAAAGAATTGTCATTATCTCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTGGACAAAAGAATGATGGACGATTGACCCGAGAGATCATGCAGTCCTGCCGGACTTG
GGAGTCAGATATCTGAGACAGCAGGCAAAGAAATTCTCAAGAGCATGGATAAAAACGGCACGATGACCATCGACTGG
AACAGTGGAGAGACTACCACTCTCCACCCCTGAAAAACATCCCCAGATCATCTTAACGGGATTCCAG
ATCTTGATGTGGGTGAGAATCTAACGGTCCCAGTGAAGTTCACAGTGGAGGAGGGAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGGTGGGGAGGGGGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCCTCCCGAGAACACATGGGCATCGTGGCTTCACTCAGATGATTGAGAAGGGGGCCAGG
TCACCTGGCCTAACGGCATCAAGCTCAGGCTCAAACGGGGAACTCAGCCCATAAATTCTGGCCTATGAGCAG
ATCAAGGCCCTGTGGTAGTCAGGAGACTCTGAGGATTCAAGAGAGGTGGCAGGGTCTTGGCAGGGCTTGGCAGGGCC
ATGCCAGCAGCATCTACCCAATGGGTCTGAAGACCCGATGGCGTGGAGACAGCCAGTACTCAGGA
ATGCTGGACTGCCAGGAGATCTGGCCAGAGAGGGGTGGCCCTCTCACAAAGGCTATGCCAACATGCTG
GGCATCATCCCTATGCCGGCATCAGCTTCACTCAGGAGACGCTCAAGAATGCCCTGCCGACTATGCACTG
AACAGCGGGACCCGGCTGTTGCTCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCAGCTAC
CCCCCTGGCCCTAGTCAGGAGGAGCTGGCAGGGCAAGGCTTATTAGGGGCTCAGGAGGTGACCATGAGCAGCCTC
TTCAACATATCTCGGAGCGAGGGGCTTCGGCTGTACAGGGGCTGGCCCAACTTCATGAAGGTATCCCA
GCTGTGAGCATCAGCTACGTGGCTACAGAGAACCTGAAGATCACCTGGCGTGCAGTCGGTGAAGGGGGAGGGC
CGCCCGGAGTGGACTCGCTGATCTGGCCGAGCCTGGGGTGGCAGGCCATCTCATTCTGTGAATGTGCCAACACT
AAGCTGTCAGGCAAGCTGTGAAACCCCTAGCAGCACCAGCAGGAGGGGGGGAGAGCTGGCAGGGCCAGGGCTT
GTCTGTCAGGCAACGGGCTCTGGTGGCTCAGGAGAACAGGACATTCTTAGGGTCCAGGGTCAAGGTCAGCAGG
CTCGGGCTCACATGTGAAGGACAGGACATTTCAGCTGCAAGCTGGCCATAGTGAGCTGGAGGCTGGAGGGGGCT
TAGTTCTCCATTTCACCCCTGCAGCAGCTGGCCACGGCCCTGCCCTGGCTCTGCCGTCATCTCCCTGTGC
CCCTTGCTGCTGCCCTGCTGAGGTAGGGAGGGCTACAGGCCACATCCACCCCTGTCCAATCCC
ATAATCCATGTAAGGGTGGCTGAGGTCACTGGCCTCCAGGGCTGACTTCCACCTACAGCATTGAGCCTAACATGGC
TGTGAAGGAAGGAAAGGATCTGGCTCAGGAGCTGGCTGAGGCTGAGCCCTGCTGATGGCTGGGGCTCTGGCAG
CTTGGAGGCTGGAGGGCTGGCTGCCCTGGCTGCACAGAACGCAAGTGTGGCTCATGGTCTGAGCT
GGCCTGGACCTGTCAGGATGGGCCACCTCAGAACCAAACACTCAGTGTCCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTGAGGGCGAAGGGCAAGCGTTGTGTTCTGGGAGGGAAAGGTGTGGAGGCTTAATTATGG
ACTGTTGGGAAAGGGATTGGCTCAGGAGCAAGGACAAGGGGACAAATGAGCAGCTTGTGCTTCCAGAGGAAGCAGG
GAGCAGGAGCTGGCTGACTGTCAGAGCTGGCTGAGCCCTGGGGTCTCTGTCACCCACAGGGGCCAGC
GGGACCGCCCCACATTCACCTGGTCACTGCTGGACCTTATTGTTGATTGATTGAAACAGAGTTATGCTT
AACTATTGTTATAGATTGTTAATTAAAGCTTGTCTTCAAGTTCTTATTTTATTGATATTATGTTATGGTT
GATGTCACCTTCCAAAGCCGCCAGTGGGATGGAGGAGGGAGAAGGGGGCTTGGCCGCTGAGTCACATCT
GTCCAGAGAAATTCTTTGGGACTGGAGGAGGAGAAAAGGCCAGAACGGCAGCAGCACCCCTGGCTCTTCCCTGGCAG
GTGGGGAGGGCTGCCCTGGCTTAAGGTTCAAGGTTGACTGGGGCTGGAGAGAGAGGGAGGAACCTCAAT
AACCTGAAAGGTGAACTCAGTTCTGGCTGCCAGGGTTCTTATTTCACTCTTCTGAATGTCAGGCAG
TGAGGTGCCCTCACTGTAATTGTTGGTGGCGGGGCTGGAGGAGGGTGGGGGCTGGCTCCGTCCTCCCTCCAGC
CTTCTGCTGCCCTGCTTAACATGCCGCCAACTGGCAGCTCACGGTTGCACCTCCATCCACCAAGAATGACCTGA
TGAGGAAATCTCAATAGGATGAAAGATCAATGCAAAATTGTTATATGAAACATATAACTGGAGTCGTCAAAAAG
CAAATTAAAGAAGGATTGGACGTAGAAGTTGTCATTAAAGCAGCTCTAATAAAGTTGTTCAAGCTGAAAAAAA
AAA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFTYRQWKQKIVQAGDKDLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLEGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGAGCTCCACTCAGCCAGTACCCAGATA
CGCTGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTGAACCTGACATCAAACTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTTGGCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTCAGAGGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTTTGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGGCGGTGTGAGGCTCCCCGATGGTCCCC
CAGCCCACAGTGGCTGGCATCCAAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCAAATAC
CAGCTTGAGCTGAACTCTGAGAAATGTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAACGAGATCAAAGGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGTT
CTCTCTTCTTGGCCATCAGCTGGCACTTCTGCCTCTAGCCCTACCTGATGCTAAATTAAT
GTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTCAC
CACCAGATATGACCTAGTTTATATTCTGGGAGGAATGAATTCAATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAT
CTATCTCAAAGACATATTAGAAGTTGGAAAATAATTCAATGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCGCTGT
CACCTGGGAGTGGAGGACAGGGATAGTGCATGTTCTGTCTGAATTAGTTATATGTG
TGTAATGTTGCTCTGAGGAAGCCCCTGAAAGTCTATCCAAACATATCCACATCTTATATTCCAC
AAATTAAAGCTGTAGTATGTACCTAAAGACGCTGCTAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGTCAAATGATTCACTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCGATTTATAAATAACTGAGCACCTTCTTTAAACAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVGLVHFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIIRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYLMNK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTAGAACATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCCCTCCGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCTGGGGTGGTTATGGGGTCTGCCCTGGAGGGCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGATGTTCCCTCTGGAACCTCCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGCCCTATGGTCAGCCACCTCCAAGTTCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGCCCTCCAAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGCTACGGCTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCATTAGCTACACAGAGCTGCAGCAAGCTCTGCTCC
AAATGGGCTACAACCTGAGCCCCAGTTACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGAGAAGGACACAGCTGTACAAGCAACATCCGGCTCAGCTCGAGGACTTCGTCA
CCATGACAGCTCTCGGATGCTATGACCCAAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGTTCTTAGAGTGAGAGAAGTATGTGGACATCTTCTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGTGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCTGCATCATAGCCACCA
AAATAGTGAGGACCAGGGCTGAGGCCACACAGATAGGGGCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGGCAGGTCTTGTAAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCGTATGCCAGTGGTAGTGTTCATCGGCCTGTTACC
GTTAGTACCTGTGTCCTCACCAAGGCCATCTGTCAAACAGAGCCATTCTCCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAAGTGGCTGGATTCTGCCACACCCATAAAT
CCTTGGTGTGTTAACCTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGCAT
CTTGGCCAGGCTCTGCCCTGCAGCTGGACCCCTCACTGCCTGCCATGCTCTGCTCGGCT
TCAGTCCTCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTAATTGCATTTTTTC
ATTGGGCCAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYPGPPNSGGQYGSGLPPGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTGTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNVDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTSGRIDVYGFSAWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCCGTGGCAGGGAGCTGCCTCCTGGGCCTGCTCCTGGCTGTCTTCATC
TCCCAGGCCTTTGCCCGAGCATCGGTGTTGGAGGAGAAAGTTCCAAAACCTCGGGACC
AACTTGCCCTAGCTCGGACAACCTCCACTGGCCCTCTAACCTGAACATCCGAGCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTCCCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGGAGGTTCTGCAGTGCAAGAGGTGGCCTCCATCGTGGGGCTGCCTGCCATG
GATTCCCTGGCCCCCTGAGGATCCTGGCAGATGATGGCTGCTGCCGTAGGACCGCCTGGGGGA
AGCGCTGCCTGAAGAACACTCTTACCTCTCCAGTGCTGCCGTGGCCCTCGCTCCGGCAGTGGCCCTT
TGCCTGGGAGTCTCTCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCACCAAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAATTCACTGGGAGCCGGGGAAAAATCCTTCCAACG
CCCTCCCTGGTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGTACCCCTGAATCCACTG
TGTCCCTGGGAGGTGGAGGCCCTGGACTGGTGGGAACGAGGCCATGCCACACCCCTGAGGGA
ATCTGGGTATCAATAATCAACCCCCAGGTACCAAGCTGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGAAATATTAATCGGTATCCAGGAG
GCAGCTGGGAAATATTCACTTACCCAGGTATCAATAACCCATTCCCTGGAGTTGCAGTG
CCTCCTGGCTCTTGGAACATCCCAGTGCTGGCTCCCTAATCCTCAAGCCCTAGGTTGCAGTG
GGGCTAGACGATAGAGGGAAACCCAACATTGGGAGTTAGAGTCCTGCTCCGCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTCACTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSQVRWPPSWGLPAMDSWPEDPWQMMAAAEDRLGEA
LPEEFLSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGPGTGWGTRPMHPPEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSGNINRYPGGSGNIHYPGINNPFPVGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGTC
TGGGCTGCCCTTGCTCCTCTTGACCCCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTCATGAGTCCAGCTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTCAACACCATGTTGTCGCAACACATGACAGCCATTGAAGCCTGTGTCCTCTGGCCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGCCCGACCCCTGTCTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAATAAAATTGGTATGCTG

FIGURE 66

MGSGLPLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGAGGTGCCCTGGCCGAGATGCGTAGGAGGGCGAGCGCGAGAACCCCC
TTCCTCGCGCTGCCAACCGCCACCCAGCCCATGGCGAACCCCGGCTGGGCTGCTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCGCTGGGCGAGCCTGGGGCAAATACAGACCACTCTGC
AAATGAGAATAGCACTGTTTGCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTGGCTGCCTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGCAGAAGCTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCTCTCCCTGCATCTGCTCCCTCATTGCTGTGACCTGGGAAA
GGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGAAGAAGGTACTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GGCAGGAATAACTAGAGAGGAACAATGGGGTTATTCAGAGGTTTGTTCCTCTTAGTTCTGTGCCCTGCTGCACCAAG
TCAAATCTCCCTCATTAAGCTGAATAATAATGGCTTGAGATATTGTCATTGTTAGATCCTAGTGTGCCAGAA
GATGAAAAAATATTGACAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTGAAGCCACAGAAAAAGA
TTTTTTTCAAAAATGTATCTATAATTCTGAGAATTGGAAAGGAAATCCTCAGTACAAAAGGCCAAAACATGAA
AACCATAAACATGATGTTATAGTTCTGACCCACCTACACTCCCAGGAGATGAACCATACACCAGTTCA
GAATGTGGAGAGAAGGGAAACATTCACTTCACCCCTGACCTCTACTTGGAAAAAAACAAAATGAATATGGACCA
CCAGGCAAACTGTGTCATGAGTGGCTCACCTCCGGTGGGAGTGTTTGAGTACAATGAAGATCAGCCTTC
TACCGTCTAAGTCAAAAAAATCGAAGCAACAAGGTGTCGCAGGTATCTCTGTTAGAAATAGAGTTATAAGTGT
CAAGGAGGAGCTGTCTAGTAGAGCATGCGAGATTGATTCTACAAACAAAAGTGTATGAAAAGATTGTCATTCTT
CCTGATAAAAGTCAAAACAGAAAAGCATCCAAATGCAAGTATTGATTGTCTTGAGATTGTAACGAA
AAAACCCATAATCAAGAAGCTCCAAGCTACAAACATAAAGTCAATTAGAAGTACATGGAGGTATTAGCAAT
TCTGAGGATTTAAAACACCATCCATGGTGACCACCCCTCCACCTGTCATGCTGAAGATCAGTCAA
AGAATTGIGCTTAGTCTGATAAGTCTGGAAGCATGGGGGTAAGGACCGCTAAATCGAATGAATCAAGCAGCA
AAACATTCTCTGCTGCGACTGTTGAAAATGGATCTGGTGGGATGGTTCACTTGTATGAGTACTGCCACTATTGTA
ATAAGCTATCAAAAAGCAGTGTGAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
ACTTCCATCTGCTCTGGAATTAAATATGCATTTCAGGTGATTGGAGAGCTACATTCCAACTCGATGGATCCGAAGTA
CTGCTGCTACTGATGGGAGGAAACACTGCAAGTCTGTGATTGATGAGTGAAACAAAGTGGGGCAATGTTCAT
TTTATTGCTTGGGAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCTTTATGTT
TCAGATGAAGCTCAGAACATGGCCTATTGATGCTTGGGCTTACATCAGGAAATACTGATCTCTCCAGAAG
TCCCTCACTCGAAAGTAAGGATTAACACTGAATAGTAATGCTGCTGATGAACGACTGTCATAATTGATAGTACA
GTGGGAAAGGACGCTTCTCTCATGGAACAGTCGCTCCAGTATTCTCTGGGATCCCAGTGGAAACA
ATAATGGAAAATTTCACAGTGGATGCAACTTCAAATGGCTATTCAGTATTCCAGGAACTGCAAAGGTGGCAG
TGGCATACAATCTCAAGCAAAGGAAACCCAGAACACATTAACTATTACAGTAACCTCTGAGCAGCAAATTCTCT
GTGCTCCAATCACAGTGAATGCTAAATGAATAAGGACGTAAACAGTTCCCAAGGCCAATGATGTTACGCAGAA
ATTCTACAAAGGATATGTAACCTGTTCTGGAGCAATGTACTGACTGCTTCAAGGATATGCAAGTACAGAAGTT
TTGGAACCTTTGGATAATGGTCAAGGCTGATGGGAGGAAACTGCAAGCTTCAAGGATATTTCAGCTGAGGTTTTACGCATA
ACAGAAAATGGCAGATATAGCTTAAAGTCCGGCTCATGGAGGACAAACACTGCAAGCTAAATTACGGCTCCA
CTGAAAAGCAGCGCTACATACCAAGGCTGGTAGTGAACGGGAAATTGAAGCAAACCCGCAAGACCTGAAATTGAT
GAGGAACTCAGACCACCTGGAGGATTTCAGCCGAACAGCATCCGGAGGTGCATTGTTGATCAGTCCAGC
CTTCCCTGCTGACCAAATACCCACCAACTCAAACTCACAGACCTGATGACTGCTTCAAGGATTAATTTCTACGCTGGTATTG
TCTGGTGATTGGGCTGTTGAAATTGTAACTTTAAGTACCCACATTGAACCTTAACGAAGAAAAAATCTTC
AAGTCAAGAGACAGTTTGATGCTCTCAAGTAAAATACTACTGATCTGTACCAAAGGAGGCAACTCCAAGGAA
AGCTTGCATTAAACCAAGAAATCTCAGAAGAAATGCAACCCACATATTGCCATTAAAGTATGATAAA
AGCAATTGACATCAAAGTATCCAACATTGCACAGTAACTTTGTTATCCCTCAAGCAAATCCTGATGACATTGAT
CCTACACCTACTCCACTCCACTCCGATAAAAGTCATAATTCTGGAGTTAATTTCTACGCTGGTATTG
TCTGGTGATTGGGCTGTTGAAATTGTAACTTTAAGTACCCACATTGAACCTTAACGAAGAAAAAATCTTC
AAGTAGACCTAGAAGAGTTTAAACAAACATGTAAGTAAAGGATATTCTGAATCTTAAATTCTACCCAT
GTGGATCATAAACTCATAAAAAATTAAAGTCTGGAAAAGGACTTGTATTAAAAAAACACTCATGGATA
TGTAAAAGTCTGCAAGTAAATTAAAGTCTGGTTGATATTGATGCAACAGTTTCTGAAAATGATATTTCAAATTGCA
ATCCTTTCTACTGATACCTGGTTGATATTGATGCAACAGTTTCTGAAAATGATATTTCAAATTGCA
GAAATTAAATCATCTGAGTAGTCAAAAATCAAGTAAGGAGGAGCAAAACCAATGCAAGTAAATTGGAAAA
AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNGFEDIVIDPSVPEDEKIEQIEDMVTTASTYLFE
ATEKRFFFKNVSILIPENWKNPQYKRPKHENKHADIVAVAPPTLPGRDEPYTKQFTECGEKEY
IHFTPDLGGKKQNEYGPPGKLHVHEWAHLRGVFDEYNEDQPFYRAKSKKIEATRCASAGISGRN
RVYKCQGGSCLSACRIDSTTKLYGKDCQFPDKVQTEKASIMFMQSIDSVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTCCCCFVFSLLKISQRIVCLVLDKGSGMMGKDRLNR
MNQAAKHFLQLQTENGWSWGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGESEVLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGTLNSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDAKSMAVLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNA
GADSFKNDGVSYRYFTAYTENGRYSLKVRAGGANTARLKLRRPLNRAAYIPGWVNGEIEANPP
RPEIDEQTTLTLEDTSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN
FDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPKSHNSGVNISTLVLSIGSVVI
VNFILESTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACTGCCCCGGGCAGGGTGA
CAACAGGTGTCATCTTTGATCTCGTGTGCGCTGCCCTCATTCAGGAAAGACGCCAAGGTAAATTGACCCA
GAGGAGCAATGATGTAGGCCACTCTTAACCTCCCTTCTGAAACCCCAGTTAGCAGGATTACTAGAGAGTGTC
ACTCAACCAGCAAGCGGCCTTCGCTTAACCTGTGGTTGGAGGGAGAACCTTTGCGGGCTCGCTCTTAGCA
GTGCTCAGAAGTGACTTGCGTGGGGAGAGAAAGGAAAGGTCCCTCTGCTGTTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGTGAACACTGGAGATTCACTGCTATTGCTGCTGCAAGATCATCTTTAAA
AGTAGAGAGCTGCTGTGTGGTTAACCTCAAGAGGCAGAACCTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCAAACGCATGCTTCTGTGGCTAGCCAGGGAAAGCCCTTCCGTGGGGCCCGCTTGAGGGATGCC
ACCGGTTCTGGACGCTGGCTGATTCTGAATGATGATGTTCGCCGGCTGCTTCGCTGGATTTCGGGGTGGT
GTTTGCTGGTCTCTGCTGTCTATCTGCTGACATGTTGGCTGCAACCCAAAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGGCAACAGCCCCACGGGGAGGGGTACCGCCCTTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCTGAAGCGCAGATGCCACAGCTCAAGGAGGAGCTGCAAGGAGAGACTGAGCAGCTCAG
AATGGGAGTACCAAGCCAGCGATGCTGCTGGCTGGCTGGACAGGGAGCCCCCAGAGAAAACCCAGGCCACCTC
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGGGTAATGCTGGCTCAAGCTGGCACAGAGTATGAGCAGCTG
CCTTTGCTAGCTTACTCTACAGAAGCTGACCGACTGCCATACCGGAGGAGAACCTGTG
AGGAAGGACAAGGGATGAGTTGGAGGGCATTGAATCAGCCTGGAGGGCTGAACAATCTGAGAGAACAGC
CCCAATCACCGTCTTACACGCCCTGTGATTTCATAGAAGGGATCTACGAAACAGAGAACAGGAGCATTGAT
GAGCTACCTCAAGGGGACCACAAACAGAATTCAAACGGCTCATCTTATTCGACCATTCAGCCCATCATGAAA
GTGAAAAATGAAAAGCTCAACATGGCCAACACGCTTATCAATGTTATCGTCCTCTAGAAAAGGGTGGACAAGTT
CCGCAGTCTATGAGAATTTCAAGGGAGATGTGATTGACAGGATGGAGACTCTCAGTGTGTTTACTTTGG
AAAGAAGAAAATGAGTCAAAGGAATACTGAAACACTTCAAAAGCTGCCAACTTCAGGAACCTTACCTCATC
CAGCTGAATGGAGAATTTCTGGGGAAAGGGAGTTGATGTTGGACCGCTCTGGAAAGGGAGCAAGCTTCTC
TTTTCTGATGTCAGCATCTACTTCACATCTGAATTCTCAATACGTGTAAGGCTGAATACAGCCAGGGAGAAG
GTATTTTATCCAGTTCTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTGGAA
CAGCAGCTGGCTAAAGAAGGAAACTGGATTGGAGAGACTTTGGGATGAGCTGTCAGTATCGGTAGAC
TTCATCAATATAGGGTTGATCTGGACATCAAAGGCTGGGGAGAGATGTGACCTTTATGCAAGTATCTC
CACGCAACTCTAGGGTACGGACGCTGTGGAGGACTCTCCACCTCTGGCATGAGAACGCTGCAATGACGAG
CTGACCCCGAGCACTACAAGATGTGATGCAGTCAAGCCATGAACGAGGCATCCCACGGCAGCTGGCATGCTG
GTGTTAGGCACGAGATAGAGGCTACCTTCGCAACAGAAACAGAAAGACAAGTAGCAAAAAACATGAACTCCAGA
GAAGGATTGTGGGAGACACTTTCTCTGGCTTGGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGACAAAAGAATTGGACTGATGGGAGATGAGAGACTGGCTGGGAGAGATGGCTTCTCTGTGGGCTTTACACAGA
AATCAAAATCTCCGTTGCTGCAAAACTAACCGGACTTGCACCCCTGTGACAAAGGCAGAATGCTTG
AGATTATAAGCTAATGGTGTGGAGTTTGATGTTGTTACAATACACTGAGACCTGTTGTTGTGCTATTGA
AATATTCTGATTTAAGAGCAGTTGTTAAAAAAATTCAATTAGCATGAAAGCAAGCATATTCTCTCATATGAATGA
GCCTATCACAGGGCTTAGTTCTAGGAATGCTAAATCAGAAGGCAGAGGAGGAGATAGGCTTATTATGATAC
AGTGAGTACATTAAGTAAATGGACAGATGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
TAACCAAAATATCTGTTATCTTTGTTGCTCTTAACTGTCCTGGTTTACTGTCCTGGTTTCTTCTTCTTCTT
CTTCTGTTTACAGAAAAGGAAACTATTCACTGGTACATGTTGCTGTTTACTGTCCTGGTTTACTGTCCTGGTT
CTCTCTGAGAAGTGGGAGCCGCTTCTACCTGTTAATAAACCAAAGTACCTGGTGAACCAAACATCTTCTC
AAAACAGGGTGCCTCTGGCTTGGCTTCCATAAGAAGAAATGGAAAAATATATATATATATATATATATG
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTGCTACATGTTATCCACCCAGGGCAGGTGGAAG
TAAGTGAATTATTTTAAATTAAGCAGTTCTACTCAATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
CAAACATTCTTAAATTAAGCTTAACATAGAGTGTGTTCTGTTCTGTTGCTACAGTAAACTCATTGTTAAAGGCTCAAGA
ATGCATGAGCTAATTATCTCTTGTGAGTCCTGCTGTTGCTACAGTAAACTCATTGTTAAAGGCTCAAGA
ATTCAAGCTGTTGGTGTGTTAAAAATGCAATTGTTGACTGGTAGTTATGAAATTAAATTAAACACAGG
CCATGAATGGAAGGTGGTATTGACAGCTAATAAAATGATTGTTGAGTAA

FIGURE 72

MMMVRRGLLAWISRVVVLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALETLNNPA
ENS PNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMAN
TLINIVPLAKRVDKFQRQFMQNFRMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFQSQY
NPGIIYGHDAVPPLQQQLVIKKETGFWRDFGFGMTQCQYRSDFINIGGFDLDIKGWGGEDVHLYR
KYLHSNLIVV RTPVRGLFHLWHEKRCMDELTP EQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTGTCTGGGATCCA
GAAACCCATGATAACCTACTGAACACCGAATCCCCTGAAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTGCCTGTCTAGTCCTAGTCCTCAAATTCCAGTCCCCTGCACCCCTTC
CTGGGACACTATGTTCTCCGCCCTCCTGCTGGAGGTGATTGGATCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGGCCAGCCTTACCCGTAGTGT
GGAAACAATGCCAGTCGCCATCGATATTACAGACAGACAGTGTGACATTGACCCGTAGTGT
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTCACCCGTATCTGGGTGACTCCCCGAAAATATGTAGCT
GCCAGCTCCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGGTCAGAACACCAAGATCAACAG
TGAAGCCACATTGCAGAGCTCCACATTGTACATTATGACTCTGATTCCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCCTCAGGGCTGGCTGTCCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCACCTGCATGAAGTCAGGCATAAAGATCAGAACACTC
AGTGCCCTCCCTCAACCTAACAGAGCTGCTCCCCAACAGCTGGGCAGTACTTCCGCTACAATG
GCTCGCTCACAACTCCCCCTGCTACCAGAGTGTGCTCTGGACAGTTTTATAGAAGGTCCAG
ATTTCATGGAACAGCTGGAAAAGCTTCAGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGCATGGTCTTGCTTCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTGGTTGGC
TGTCTCTGCCCTCCTGGCTGTTATTGCTAGAAAGATTCCAAGAAGAGGCTGGAAAA
CCGAAAGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCATAAAATTCCCTCAGATAC
CATGGATGTGGATGACTTCCCTCATGCCTATCAGGAAGCCTCTAAATGGGTGTAGGATCTGG
CCAGAACACTGTAGGAGTAGTAAGCAGATGTCCTCCCTGGACATCTTAGAGAGGAAT
GGACCCAGGCTGTCACTCCAGGAAGAACTGCAGAGCCTCAGCCTCTCCAAACATGTAGGAGGAA
ATGAGGAAATCGCTGTGTTAATGCAGAGANCAAACCTCTGTTAGTGCAGGGAAAGTTGGG
ATATACCCCAAAGTCTCTACCCCTCACTTTATGGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTAGGATAAAAGAGTTGCTGTTGAAGTTGTATATTGATCAATATATTGAAATTAAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQS PIDIQTDSVTFPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCCTGGGACGGCAGTCCCTGTGTC
TCTGGTGGTTGCCTAACCTGCAAACATCACCTCTTATCCATCAACATGAAGAATGTCCCTACA
ATGGACTCCACCAGAGGGCTTCAGGAGTTAAAGTTACTACACTGTGCAGTATTTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGCTGACAGCTCC
AGAGAAGTGGAAAGAGAAATCCAGAAGACCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGACCAACCAC
ACGCTGGTGTCACTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTGCCGTGCTCAGCCTCTGAGAACAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTCAAGGCTAAATCATCTCTGGTATGTTGCCATATCTATTACCGTGTGTTCTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTGGCAAAGAGAAACACCCAGCAAATT
GATTTGATTTATGGAAATGAATTGACAAAGATTCTTGTGCCTGCTGAAAAAAATCGTGATTA
ACTTTATCACCCCTCAATATCTCGGATGATTCTAAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGTATCCAGCCTTAATGATCCTCAGCCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGGTATGCTTCGCAATTGATGAAATTGGACTCTGAAG
AAAACACGGAAGGTACTTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACACTGACATTGTCGGGGCCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTGGAGTCGCAAGCAGCGTTGCAG
TCCTGGGCCGCAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCCTGGTCACTGGGA
TCCCCAAACTGGCAGGCTGTATTCTCTGCTGTCCAGCTCGACCAGGATTAGAGGGCTGCG
AGCCTTCTGAGGGGATGGGCTCGGAGAGGAGGGCTTCTATCTAGACTCTAGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAATGAAACCTATCTCATGCAATTGAGGAATGGGGTTATA
TGTGCAAGTGGAAAACTGATGCCAACACTTCTTTGCCATTGTTCTGTGCAAACAAAGTGAG
TCACCCCTTGATCCCAGCATAAAAGTACCTGGATGAAAGAAGTTTTCCAGTTGTCACTGT
CTGTGAGAATTACTTATTCTCTATTCTCATAGCACGTGTGATTGGTCATGCATGTA
GGTCTCTAACATGATGGTGGGCCTCTGGAGTCCAGGGCTGGCCGGTTGTCTATGCAGAGAA
AGCAGTCATAAAATGTTGCCAGACTGGGTGCAGAATTATTAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGYNEFDKRFFVPAEK
IVINFITLNISDDSKISHQDMSSLGKSSDVSSLNDPQPSGNLRPPQEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTDICAGPEEQELSLOQEEVSTQGTLLESQA
ALAVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSFDQDS
EGCEPSEGDSLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMAN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACCTGCTGCCCTCTGACAC
CTGGGAAGATGGCGGGCCCCTGGACCTTCACCCCTCTCTGTGGTTGCTGGCAGCCACCTTGATC
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATTCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGTCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGCCAATGACCA
GGAGCTGCTAGTCAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
CCCACCCGCGCTGGTCCCTCAGTGAUTGTGCCACCAGCCATGGGAGCCTGCGCATCCAACCTGCTGA
TAAGCTCTCCCTGGTGAACGCCCTAGCTAACAGGTCAACCTCTAGTGCATCCATGGTCAAGACCA
CCAATCTAGTAAAAACCAAGCTGTGTCCTGATCGAGGCTTCCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTACACCATTCAAGCTCTACCTGGGGCCAAGTGTGGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCAGCTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGAGCTGCAGTGGCTGCTGTCTCCAGAAGA
ATTCAATGGTCCCTGGTGGACTCTGTGCTTCCTGAGAGTGGCCATCGGCTGAAGTCAAGCATGGG
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTATAGACCAAGGCCATGCCAAGGTGGCCAACTGATCGTCTGGAGTGTT
TCCCTCCAGTGAAGGCCCTCCGCCCTTGTCACCCGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTACCAACTTAACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACCTGGGATTGGCTGGTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAAGATCTGGGGTCCAGTGTGATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCCCTACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAAGGCTGG
GTCCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLSAMREK
PAGGIPVLGSLVNTVLKHI IWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPTLDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPEAHRLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGCTACATCCTAGGCCTTCTGGGCTTTGGGCACACTGGTGCCAT
GCTGCTCCCAGCTGAAAACAAGTTCTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTCTGGCCTGCCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGGGCATGAGATGCACAGTCTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGTAGCAGGTGGAGTCTTTTCATCCTGGAGGCCCTGGGATTC
ATTCCCTGCTGGAATCTCATGGGATCCTACGGGACTCTACTCACCACACTGGTGCTGACAG
CATGAAATTGAGATTGGAGAGGCTCTTACTGGGCATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTCTGCTCATCCCAGAGAAATCGCTCCAACACTACGATGCCCTAC
CAAGCCCAACCTCTGCCACAAGGAGCTCTCAAGGCCTGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAGAACCAAGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCGAGGGCACAGGTGAGGGACACTACCAACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTGAATTGCAAGGATGCTGCCATGCCAGCCTTCTGTTTCC
TCACCTTGCTGCTCCCTGCCCTAAGTCCCCAACCTCAACTTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGCCTCTGGTTACCTGGACTCCATCCCCAACCCACTAATCACA
TCCCACGTACTGACCCCTGTGATCAAAGACCCCTCTCTGGCTGAGGTTGGCTCTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTGTGGCATTGCTCTAACCTACTCTCAAGCTTC
CCTCCAAAGAAACTGATTGGCCCTGGAACCTCCATCCACTCTTGTATGACTCCACAGTGTCCA
GACTAATTGTGCTGACGAAATAAAACCATCCTACGGTATCCAGGGACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSAAISSLACIIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACCGGTCCGCCTCTCCCTCTGCTGGACCTCCCTCGTCTCCATCTCTCCCTCTTC
CCCGCGTTCTTTCCACCTTCTTCTTCCACCTTAGACCTCCCTGCCCTCCTTCC
GCCCACCGCTGCTTCCCTGGCCCTCTCGACCCCGCTAGCAGCAGACCTCCCTGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCTCCGTGTCCTTCGTCTCCCTCCGACTCCGCTCCGG
ACCAGCGGCCTGACCCCTGGGAAAGGAGATGGTCCCGAGGTGAGGGTCTCTCCTGCTGGGA
CTCGCGTCTGTTCCCCCTGGACTCCCACGCTGAGCCGCCAGACATGTTCTGCCTTT
CCATGGGAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCCTGCGCTGTACCTGCTCAGAGGGGCCCATGTGAGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCGACGCTGTGACGGAGCCACAGCAATGCTGCTCCAAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACCAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
AGATCTCAGTGCCTATGAGCTGTTCCCTCCGCCCTGCCAACCCAGTGTGTCTGAGCTGC
ACAGAGGGCCAGATCTACTGCGGCCTCACAAACCTGCCCGAACCGAGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGTGCTGGAGAAAG
AGAGGCCGGCACCCAGCCCCACTGCCCTAGGCCCTCTGAGCTCATCCCTGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACGTCAAGATGCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGGAAAGACGTACTCCCACGGGGAGGTGTGGCACCCGCCCTCCGTGCCCTCGGC
CCCTGCCCTGCATCCTATGCACCTGTGAGGATGGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTGCCAGAGG
ACAAAGCAGACCCCTGGCCACAGTGAGATCAGTTCTACCAAGGTGTCCAAGGCACCGGGCCGGGTC
CTCGTCCACACATCGGTATCCCAAGCCCAGACAACCTGCGTCGCTTGCCCTGGAACACGAGG
CTCGGACTGGTGGAGATCTACCTCTGGAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACACAGCCAGAATCTTCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTCCAGAAAGAGGCACAGCACTCCGACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTCCTAGCCCAGACCCCTGGAGCTGAAGTCACGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAACAGTGCAGATATGAGCTGTATAATTGTTATTATATTAATAAAA
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKGPGTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLSDQESQEARNPERGTALPTARWPDRRSLERLPSPDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCCCTCTCTAATCCATCCGTACCTCTCTGTCA
TCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTGGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGGCCAGAACAGCCTGTCCAGGCCTGGTGGGGAG
GACGCAGCATTCTCTGTTCCTGTCTCTAACAGCCAATGCAAGGCCATGGAAGTCCGGTTCTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAATGGTAAGGATTCTATTGCGGAGGGCGCATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGCCCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAACGGCATCTGGAGCT
ACAGGTGTCAGCACTGGCCTCAGTTCCATCAGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCACAAGGACAGGATTTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGCCGTGTTGATGTGGAGATCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTCGAGCCTATATCGTGGCACCTGGTACCAAAGTACTGGAAACTCTGCTGTGGCCTA
TTTTTGGCATTGTTGACTGAAGATTCTCTCAAATTCCAGTGGAAATCCAGGGGAACTGGACTG
GAGAAGAAAAGCACGGACAGGCAGAATTGAGAGACGCCGGAAACACGCACTGGAGGTGACTCTGGATCCAG
AGACGGCTACCCGAAGCTCTGCGTTCTGATCTGAAAAGTGTAAACCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAGGTGGCGTGGAGTGTGCCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTGTCTCCGATCATGGTACTGGTCTCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAAATCCCGTTTATCAGCGTCTCCCCAGGACCCCACCTACAAAAAATAGGGCTTCTGGACTA
TGAGTGTGGGACCATCCTCTTCAACATAATGACCACTGGCTTCTCAGAGTTCCAAGCAGGGAAACATTA
AAGGCTTATTGAGGCCCTACATTGAGTATCCGCTCTAATGAGCAAATGGAACATCCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAACAGGCCCTCTGGCAAAGGGCTCTGCAATCCCAGAGACAAGCACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTCTCCCCAGGGTGAATGTAGATGAATCACATCCCACAT
TCTTCTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCAAGGTGGCTTCCA
GATGAAGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGAGGGAAAGG
CTGACATTACATTAGTTGCTCTCACTCCATCTGCTAAGTGATCTTGAATACCACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCTAATGGTTGTTCAATTATATTACACTTCAAGTAAAAAAA

FIGURE 84

MALMLSIVSLLKLGSGQWQVFGPDKPVQALVGEDAASFCLSPKTNAEAMEVRFTRGQFSSVVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVPPLISITGYVDRDIQLLCQSSGWFFRPTAKWKGPGQGDLSTDRTNRDMHGLFDVEISL
TVQENAGSISCSMRRAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAEQELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVRLNGEHLYFT
LNPRFISVFPPRTPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTAACCCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGGGCGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTCTCCTACCCCTCGCATGGCTGGATT
TACCCGGCCCAAGTAGTCATGGCTACTGGTTCCGGAAAGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTGGGAGTGTGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGTGCGGGGAGA
TACCTCTTCGTATGGAGAAGGAAGTATAAAATGAAATTATAAACATACCGGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCACCACCCGCTCCTCGGTGCTCACCCCTATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCCTACCCGCTCAGAACTTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTATCTCTGTCACTCCCAGAGGGCAGTCT
GCGCCTGGTCTGTGCAAGTGATGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCCTGTGCCCTCACAGCCCTCAAACCCGGGGTGCTGGAGCTGCCCTGGTGCAC
CTGAGGGATGCAGCTGAATTCAACCTGCAAGAGCTCAGAACCCCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGAGCTGGAG
CCACAGCCCTGGTCTTCCTGCGTCATCTCGTTGAGCTGAGGCTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGAGATACTGGCATAGAGGATGCAAACGCTGTCAGGGTTC
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGAGAAGACAGTCCCCAGACCAGCCTCCCCAG
CTTCTGCCCGCTCCCTCAGTGGGGAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGT
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCCGTATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTACAGCCCTAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTAAAAGCACACAGGCTTGTAGAGTCAAAGTATCTAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTAACTAAAGACAGACAAATTCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTQEGLCVHPCFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSVPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLGNSSL
SLPEGQSLRLVCAVDADVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSSARPAAGVGDTGIE
DANAVRGASQGPLTEPWAEDSPPDOPPPASARSSVGEGELOQYASLSFOMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAGGAATGGACCTGT
TCTTCGTCTCCATCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGA
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCTGGTGGCCAGCGTGATGAGAATGACATGCGTGGAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGG
CAACTACAACACCTTGGATCTGCAGAGGCGGCCAGGAGCGATGACTACAAGAACCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAAATATCCAGTGAAAATATGGAGAAGGAAAGTGTGGACTGACAACGCC
CGGTGATCCCTGTGGTCTATGATTTGGCAGGCCAGAAAACAGCATCTTATTACTCACCCAT
GGCCAGCGGGAAATTCACTGCAGGATTTGTTCAAGGTATTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTCCAGAGGCCAGTCCCAGCAGTGTGGAGATTTCTGGTTGATTGGAGTGGATAT
GGAACTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAAACCCAGACCTCTCCCAACCATGAGATCCAAAGGATGGAGAA
CAACTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRSSQQGSKADYPEDGNWANYNTFGSAAAT
SDDYKNPGYYDIQAKDLGIWHVPNPKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTNDGPVIPVYDFGDAQKTASYYSPYQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGCAGGGAGACTTCAGGAGTCGCTGTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTGTCCCCGAGGGCATGGGCCGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTGTATCACCTTATTCTGGTCCCAGGACAGCAACATAACAGGCCTG
CCTGCCCTCTCACGTTCACCCCCGAGGAGTATGACAAGCAGGACATTCAAGCTGGTGGCCGCGCTCT
CTGTCACCCCTGGGCCTTTGCAGTGGAGCTGGCCGGTTCCCTCTCAGGAGTCTCCATGTTAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCATCCGTGGCCCTGCCCTTTCAT
ATTCAAGCAGTGGGAGTGCACTACGTATTGGTACATTTGTCTCTGCAGTGCCCTCCAGCTG
TCACTGAAATGGCTTATTCGTACCGTCTTGGCTGAAAAAGAAACCTCTTGATTACCTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCGCTTCGTATTCTGGAAAGAAGGAAG
GCATAGGCTTCGGTTTCCCCTCGGAAACTGCTCTGCTGGAGGATATGTGTTGAAATAATTACG
TCTTGAGTCTGGATTATCCGATTGTATTAGTGCCTTGTAAATAATGTTGTAGTAACA
TTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC SALPAVTEMALFV
TVFGLKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

Copyright © 2007 by the American Society for Biochemistry and Molecular Biology, Inc.

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAAGAATGCAACTGAC
TCGCTGCTGCTCGTGTTCCTGGTGCAGGGTAGCCTATCTGGTATCTGTGCCAGGATGATG
GTCCTCCGGCTCAGAGGACCTGAGCGTGATGACCACGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAACGAGTCCCAGGGCATGGCAATTCAACTCTCCAGGGCTGCT
GGCCCCGCTGGGAGGCTTGGGCATTCTGGGCAGCCCCCAACGCCGAACCACAGCCCC
CACCCCTCAGCCAAGGTGAAGAAAATCTTGCTGGGCAGTTCTACTCAAACATCAAGACGGTG
GCCCTGAACCTGCTCGTACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTCCCCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTCATGAAAGCCAAGGCCTCCAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGCCGCCGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACATAC
CATAGTGATAACCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCCGGCAGGGAAAGGGTTGGCCTC
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCAAGTGGGCCAGGGCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGTGGTCCAAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCTGTGGCAGGCCATCGAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAC
CCTTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGAGATTTCATCAGTGTGGACA
GCCGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTTAGGAGCCAGTCAGCAGGGTGGGTGGG
GCCAGAGGAGCTCCAGCCCTGCCTAGTGGCGCCCTGAGCCCCCTGTCGTGTGCTGAGCAG
CATGAGGCTGAAGTGGCAACCTGGGTCTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCCAAAATCCCTCTCTGCCAGTACTCCCCCTGTACCCACCTGCTGATG
GCACACCCATCCTAACAGCTAACAGACAGGACGATTGTGGCTCCACACTAACGGCCACAGCCCATC
CGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTGGAGCATCCATGTCCCG
GAGAGGGTCCCTAACAGTCAGCCTCACCTGTCAAGACGGGTTCTCCGGATCTGGATGGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTC
TGTGTGTCTGTCTGTGGTGGGGAGGGAGGGAAAGTCTTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGAGCAGGAATAAGCTTGGCCGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPKVVCVYIAFYSTDYRLVQKCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCC**ATG**ACTGCGGCCGTGTTCTCGGCTGCGCCTTCATTGCCTCGGGCCTGCGCTGCC
CTTTATGTCTTCACCATGCCATCGAGCCGTTGCGTATCATCTTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATTGTCGCCCTGTTGGTCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGAGCGTTGTCTGTATATCCAAGAA
ATGTTCCGATTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTCATGCGACTGCTGGCCTATGTTCTGGCTGGGCTTGGAAATCATGA
GTGGAGTATTTCCCTTGTGAATACCCCTATCTGACTCCTGGGCCAGGCACAGTGGGCATT
GGAGATTCTCCTCAATTCTCCCTTATTCAAGCTTCATGACGCTGGCATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTGATGGCTGTGAGAAGAAAAGTGGGCATCCTCTTATCGTC
TCCTGACCCACCTGCTGGTGTCAAGCCCAGACCTTCATAAGTTCTTATTATGAAATAACCTGGCG
TCAGCATTATAATCCTGGTGCATGGCACCTGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTGCCAAGACAAGAACTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAAGCACTCCCAAACCGCAGACTACATCTTAGAGGAAGCACAAC**TGTGCCT**
TTTCTGAAAATCCCTTTCTGGTGAATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGP GTVGIHGDSPQFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKWGILLIVLLT
HLLVSAQTFISSLYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKNFLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTTGAGAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAACATGAGGAGTACGATCTACTGTTTGTCTTAGGATCAC
TCGGTCATTACCACAGCTAACCTGTTGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTTAAAGTCTGATACCATTAA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAATCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTCACACAACTTGGAGCCCAGGGCACTATCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTGTTCCGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCAGCAGGCCCTCCCAACTCCAGTGGCACAGATGACGACTTGCAGTGACCACCC
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTCAAATTTCAACTAAGCTGCCTCGAATTGGTGATACATGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTAGAAGAAATTAAATTCTTAAATT
TACCTGAAAATATTCTGAAATTCAGAAAATATGTTATGTAGAGAATCCAACTTAAAAAA
CAATAATTCAATGGATAAAATCTGTCTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTGGAAAACTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRPLPTSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTTGTCTGGCAGCTGGACCAAGGGAGCCAGTCTGGGCGCTGGAGGGCTGCTCTGACCATG
GTCCTGCCTGGCTGCTTGTGTCCTGCCCTCCCCCAGGCTCTCCCAAGGCCCAGCCTGC
AGAGCTGTCTGTGAAAGTTCCAGAAAATCTGGGAAATTCCCTTATACTGACCAAGTTGC
CGCTGCCCGTGAGGGGCTGAAGGCCAGATCGTCTGTCAGGGACTCAGGCAAGGCCAGTGC
GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGCTGGTACCCAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACCAAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTGTGGGTCAC
AGCCTGTGCTGTGACGTGAAGGATGAGAATGACCAGGTGCCCTTCTGAGGGCTCAGACCG
AGAGCTGGCTGAGCCGGGTACCAAGGCCCTGGCATCCCCCTCCCTTCTGAGGGCTCAGACCG
GGATGAGCCAGGCACAGCCAACCTGGATCTCGATTCCACATCCTGAGGCCAGGCTCAGCCAGC
CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGCTCTGGCCCTCAGCCCAGGGAGC
ACCAGCCTGACCAAGGCCCTGGAGAGGACCTACCAGCTGTGGTACAGGTCAAGGACATGGTGA
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAGGTCTCCATAGAGACACTGGGTG
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTACACCGCACCACATGCCCAAGGTA
CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCCCTTGAAGTGA
TGCAGAGGGAAACCTCTACGTGACCAAGAGAGCTGGACAGAGAAGGCCAGGCTGAGTACCTGCTCC
AGGTGGGGCTCAGAAATTCCATGGCAGGGACTATGCGGCCCCCTGGAGCTGCACGTGCTGGT
ATGGATGAGAATGACAACGTGCTATCTGCCCTCCCCGTACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGAAGTAGACTGTCACTGAGAGGATGCAGATGCCCGGCTCCCCA
ATTCCCACGTTGTGATCAGCTCTGAGCCCTGAGCCTGAGGATGGGTAGAGGGAGAGCCTTC
CAGGTGGACCCCACCTCAGGCAGTGTGACGCTGGGGCTCCACTCGAGCAGGCCAGAACAT
CCTGCTCTGGTGTGCCATGGACCTGGCAGGGCAGAGGGTGGCTCAGCAGCAGTGTGAAG
TCGAAGTCGCACTGAGGATCTGGAGGCCGGACTCTGGTGGCCATGCTAACGCCATTGATGCTGA
ATAAGCCTCCCTGAGGATCTGGAGGCCGGACTCTGGTGGCCATGCTAACGCCATTGATGCTGA
CCTCGAGCCCGCCTCCGCTCATGGATTGGCCATTGAGAGGGGACAGAAGGGACTTTG
GCCTGGATTGGGAGCCAGACTCTGGGATCTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCATGAGGTGGTGGTGGCTCAGAGTGTGGCAAGCTGGTGGGGCCAGGCC
AGGCCCTGGAGCCACCGCCACGGTGAAGTGTGGCTAGTGGAGAGAGTGTGGCAAGCTGGTGGGG
ACCAGGAGAGCTACGAGGCCAGTGTGCCCCATCAGTGGCTCCACTCAGGTTCTCCCTAGTC
CAGGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTTCTCCCTAGTCATGACTCAGAGGGCTGG
CTGCATTGAGAAATTCTCCGGGAGGTGCAACACGCCAGTCAGCTGCAGGGGCCAGCCTGGGG
ACACCTACACGGTGTGGTGGAGGCCAGGATACAGCCCTGACTCTTGCCTGTGCCCTCCAA
TACCTCTGCACACCCCGCAAGACCATGGCTGATGTGAGTGGACCCAGCAAGGACCCGATCT
GGCCAGTGGGACGGTCCCTACAGCTTACCCCTGGTCCCAACCCACGGTGCACAGGGATTGG
GCCTCCAGACTCTCAATGGTCCCATGCCTACCTCACCTGGCCCTGCATTGGTGGAGGCCAGT
GAACACATAATCCCCGTGGTGGCTAGCCACAATGCCAGAATGTGGCAGCTCTGGTTCGAGTGAT
CGTGTGCTGCAACGTGGAGGGCAGTGCATGCCAAGGTGGCCCATGAAGGGCATGCCCA
CGAAGCTGTCGGCAGTGGCATCTTGTAGGCACCCCTGTTAGCAATAGGAATCTCCTCATCCTC
ATTTTACCCACTGGACCATGTCAAGGAAGAAGGCCGATCAACAGCAGACAGCGTGCCCT
GAAGGCAGCTGCTGAATGGCCAGGAGCTAGCTGGAGCTGGCCTCTGGCTCCATCTGAG
TCCCTGGAGAGAGGCCAGCACCAAGATCCAGCAGGGACAGGACAGAGTAGAAGGCCCTCCA
TCTGCCCTGGGGTGGAGGCACCATCACCACGAGGATGTCTGCAGAGCCTGGACACCAACTT
TATGGACTGCCATGGAGTGTCCAATGTCAGGGTGTGCTTGCCTAATAAGGCCAGAGAA
CTGGGCTGGCCCTATGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEPVENYGGNFPLYLTKLPLPREGAEGQIVLSDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHMAQ
VHWSSGGDVHYHLESHPGPFEVNAEGNLYVTRELDRQAQEYLLQVRAQNSHGEDYAAPPLELHVL
VMDENDNVPICPPRDPPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVGRA
FQVDPTSGSVTLGVPLLAGONILLVLAMDLAGAEGGSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDALEPAFRIMDFAIERGDTEGTFGLDWEPSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLGVPGPGPGATATVTLVERVMPPPQKLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLICIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTIAVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTLALHWEP
REHIIPVVVSHNAQMWMQOLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTVAIGIFLI
LIFTHWTMSRKKDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGTACATTGCCTGGAGGAAGGCCAAGGATCCAGCTGCCAACGCC
AGTCCAAGATTCTTCCCAGGAACACAAACGTTAGGAGACCCACGCTCTGGAAAGCACCAGCCTTA
TCTCTCACCTTCAGTCCCCTTCTCAAGAACTCTCTGTTCTTGCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTTGCTTCCAGCACAAGAGACAGATGAAGATGCAGAAAGGAATG
TTCTCCTTATGTTGGTCTACTATTGCATTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCAGTGGAGCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTGCACCTCCAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCAGTGGGCCAGCACAGTCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCAGTGCACCAACTCTGAGTCCAGCAGTCAGTGTCC
AGTAGGGCCAGCAGTGCACCAACTCTGAGTCTAGCAGCAGTCTCCAGTGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACCTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTCAGTGTCCAGTGGGCCAGCA
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCAGCAGCTCCAGTGGGCCA
GCACAGCCACCAACTCTGACTCCAGCAGCAGTGTCCAGTGGGCCAGCACAGCCACCAACTCTGAG
TCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCAGCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCAGCACACTCTGAGTCCAGCACACCCCTCC
AGTGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCAGTGTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACA
CCTCCAGTGGGTCAAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGCCAGCAC
GCCACCAACTCTGACTCCAGCACAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCAGCACACTCTGAGTCCAGCACAAACCTCCAGTGGGCCA
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAAACAGCAGCTGTGACTGGA
ATGCACACAAACTCTCCATAGTCATCTACTGCAGTGAAGTGGCAAAAGCTGGTGGGTCCCTGGT
GCCGTGGAAATCTCCTCATCACCCTGGTCTGGTTGTGGCGCGTGGGGCTCTTGCTGGGC
TCTTCTCTGTGTGAGAAACAGCCTGCCCCCTGGAGGGAAATCATGGAGCCCCACAGGCCAG
CATGGCTCAACCATGGCTTGGAGGAGACCAAGTATCATGAGTGCATGGAGATGAGCAGGAGGA
GTGGAGCTCTAACTGGTTGGAGGAGACCAAGTATCATGAGTGCATGGAGATGAGCAGGAGGA
ACAGCGGGCCTGAGCAGCCCCGGAAAGCAAGTGCCTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCTCTTCAATTCCATCCCAGGAGACCCCTCCAGCTTGTGAGATCTGAA
AATCTTGAAGAAGGTATTCTCTACCTTCTGCTTACAGACACTGGAAAGAGAAATACTATAT
TGCTCATTTAGCTAAGAAATAATACATCTCATCTAACACACAGCACAAAGAGAAAGCTGTGCTTG
CCCCGGGTGGGTATCTAGCTGAGATGAACCTCAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEATNSNETSTSANTGSSVISSGASTATNGSSVTSGVSTATISGS
SVTSNGSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNGSSVTSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTTSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNLSLRNTFNTAVYHPHGLNHGLGPGGNHGAPHRPRWSPNWFWRPVS
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCCTCCCGTTACGGGATGAATTAAACGGCGGGTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCCACGCACCCCACTCGCGTCGCCGCGCGTGCCTGCTGTACAGGTG
GGAGGCTGGAACTATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGAAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAGCTTGGAAATCAT
GGTGTCATGGAAGGGATTACTTACTGACTCTGTTGGGAAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAA
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTGGTGTAAAGT
GATTATAACTGGGATGCATTGTCCTGGAGAAAGAAGTGTCAATTATCATGAACCACCGAACAA
GAATGGACTGGATGTTCTGTGGAATTGCGTGTGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGCTCAAAGGTGTTGGATTGGTGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCAAGGAAATGGAAGGATGACAAGAGCATTGCAAGACATGATTGATTACTTTGT
ATATTCACGAACCACCTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTGTGGTAGACCGTCAAGAGAAGGTAAGAACCTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTT
CCCAGGGAAATCCACTTACGTCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGGCACAAACGGTGGGAAGAGAAGAAGAGGAGGCTGCCTGCCTATCAAG
GGGAGAAGAATTTCATTACCGGACAGAGTGTCAATTCCACCTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCGTGTGCAAGTGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGGGTATTATAATCACCATTGTAATCTTGTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACATTGCAATTGCAACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAATGAGTAAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATATTG
GAAATGTTCTAAACCTTCAAGCTCAGATGCATTGCACTGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAATTGCTGGGAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGCAACTGTGTACATAGCAGGGAGTGATGGGGTGAATAACTT
GGGCCAGAATATTAAACAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTPVALLETMFGVKVIITGDAFVGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLIKEICLKASLKGVPFGFWAMQAAAYIFIHRKWDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDNAVHDITVA
YPHNIPOSEKHLLOQDFPREIHFHVRYPIDTLPLTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCTGCGCCTGAGACAGCTGGCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCACTGTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTGGTTCTCAGTTCTACGAGCTGGTGTCAAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCCAGGCCCTGGGGGGAGGACGCCGTGTCCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCATGGAAGTGCAGGTTCTCAGGAATCAGTCCATGCTGGTC
CACCTCTACAGAGATGGGGAAAGACTGGGAATCTAACGAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTAAGGCTAAAAAACATCACTCCCTCGG
ACATCGCCTGTATGGGTGCTGGTTCAAGTCCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACCTCTCATTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGCTCAGGGTGGTCCCCCAGGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTCAGACTCCAGAGCAAATGCCAGATGGGTACAGCCTGTATGATGTTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGACGAGATCATGA
GGTGGAAATCCAAGGTATTGATAGGAGAGACGTTTCCAGGCCCTCACCTGGCCCTGGCTTCTA
TTTACTCGGGTTACTCTGTTGCTGTGTTGATGGGATGATAATTGTTCTC
AAATCCAAAGGGAAAATCCAGGGCGAACTGGACTGGAGAAGAAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTACCCGAAGCTCTGCG
TTTCTGATCTGAAAATCTGAAACCATAGAAAAGCTCCCAGGAGGTGCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAATGTAGGGTGGTATGTTGGAGTGTGCGGATGACGTAGACAGGGGAAGAAC
ATGTGACTTGTCTCCAAACATGGTATTGGTCTCAGACTGACAACAGAACATTGTT
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCTCACCGAGTAGGGTCTT
GGACTATGAGGGTGGACCCTCCTCTCAATACAAATGACCAAGTCCCTTATTATACCCCTGC
TGACATGTCAGTTGAAGGCTTGTGAGACCCCTATATCCAGCATGGCATGATGACGAGGAAAG
GGGACTCCCATTCTCATATGTCAGTGTCTGGGATTGACAGAGAAGACCCCTGCTTAAAGGGC
CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCTCCAGTGGGAGCTTCTGCCCC
CCGGAGCCTGCGCACAGAGACTCACGCCCTACTCTCTTAGGGAGCTGAGGTTCTGCCCC
TGAGGCCCTGAGCAGCAGTCAAGCTTCCAGATGAGGGGAGTGGCTGACCCCTGTGGGAG
TCAGAGCCATGGCTGCCCTGAAGTGGGGAGGAATAGACTCACATTAGGTTAGTTGTGAAAAA
CTCCATCCAGCTAACGATCTTGAACAAGTCACCAACCTCCCAGGCTCCTCATTGCTAGTCACGG
ACAGTGATTCTGCCCTCACAGGTGAAGATTAAAGAGACAAACGAATGTGAATCATGCTTGCAGGTT
TGAGGGCACAGTGTGCTAATGATGTTTATATTACATTTCACCCATAAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTCTATACCAATCACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTAAAGAATAAAGAGGGAGTAGGATTTCACTGATTCTATAAGCCACGCAT
TACCTGATACCAAAACCCAGGCAAGAAAAGAGAAGAGGAAAGAAACTACAGGTCCATATCC
CTCATTAAACACAGACACAAAAATTCTAAATAAATTAAACAAATTAAACTAAACAATATTTA
AAGATGATATATAACTACTCAGTGIGGTTGTCCACAAATGCAGAGTTGGTTAATATTTAAAT
ATCAACCAAGTGTAAATTCAAGCACATTAATAAGTAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFQPTAKWKGPGQGDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCALCGVVMGMIIIVFFK
SKGKIQAELDWRRKHGQAEILDARKHAVEVTLDPETAHPKLCVSDLKTVTTHRKAQEVPHSEKRF
TRKSVVVASQGFQAGRHYWEVDVGQNVGVYVGVCRDDVDRGKNNVTLSPNNGYWVRLTTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGAAACCCTGGTTATCGCCTCGTCATCTTCATATCCCTGATTGTCTGCCAGTGTGCATGG
CTCACTGTTCAATTATGTGAGATATAATCAAAGAACCTACAATTACTATAGCACATTGTCATTTACAAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGCTTAACAATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTAAAAATGCATTTATAAATCTCCATTAGGGAAGAAATTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAACATGGAGTGTGGCTCATATGCTGTTGATTGAGATTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTGTGTTACATGAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAAATTAAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCCGA
ACACGAAGAAGTAAAACCTCTAGGTCAAGAGTCTCAGGATGGGATGGAGTCATCGCTGAGAACCTTAATTAA
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGAGAACCTTAATTAA
TTGTGAGTGCTGCTCACTGTTACAACATATAAGAACCTGCCAGATGGACTGCTCCCTGGAGTAACA
ATAAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAAATAATTGTCCATGAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGCAAGAGCTTCTAGCCCTGTTCCCTACACAAATGCACTGACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGTGATGTGATGTTGTCAGGAGTTGGAGCACTGAAAAT
GATGGTTACAGTCAAAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAACATTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGCAT
GCCAGGGTACTCTGGAGGACCACGGTTAGTTCACTGAGTGTGACTAGAGATATCTGGTACCTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAACACCAACAAGCCTGGTTATACTAGAGTTACGGCCTGCGGGACTG
GATTACTTCAAAACTGGTATCTAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTGTTTTG
GGTGTGGAGGCCATTAGAGATACAGAATTGGAGAAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAACTGTTGCTTGATGCATGTTCTCAGTGTGAGCAATAGTGAACACTTATGTACATAGAGAAATAGATA
GATCAACTCTGTCATCTGTGAGCAATAGTGAACACTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCTATTGTTCTAGAAGTTTGTCAAGATTGACTTGTGACATAAATTGTAAT
GCATATATAACAATTGAAGCACTCTTCTCAGTTCTCAGTGTGACTCTCATTTCAAGCAAATATCATT
TCAAGGTGCAAGAACAGGAGTGAAGAAAATATAAGAAGAAAAAAATCCCTACATTATTGGCACAGAA
AAGTATTAGGTGTTCTTAGTGAATATTAGAAATGATCATATTCAAGGAAAGTCAAGCAAAGACA
GCAGAAACCAATCACTCATTTAGAAGTATGGAAACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTTCATTTCAAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTGTGACCT
ATAATAATTATAACAACATGCAATGACTTGTCAAGCAAATTAAAGCAAATATTATTAAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAAAACACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSRQLESVKNAFYKSPLREFVKSQVIKFQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTEDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKGL
RRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTCTNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:
amino acids 21-40 (type II)

FIGURE 107

AGAGAAAAGCGCTCCAGCTGAAGCCAATGCAGCCCCCGGCTCCCGCAAGAAGTCCCTG
CCCCGATGAGCCCCCGCGTCCGACTATCCCCAGGGGGGTGGGGCACCGGGCCAGC
GCCGACGATCGCTGCCGTTGCCCTGGGAGTAGGATGTGGTAAAGGATGGGCTTCCTT
ACGGGGCTCACAATGGCCAGAGAACGATTCGTGAAGTGTCTGCCTGCTACGCCCTCAA
TCTGCTCTTGTTGTAATGTCATCAGTGTGTCAGTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTAACTGCAGAACGAGGGTAGAGGAAGCAGTCATTGACTTACCTCCT
GTGGTTCATCCGGTCTGATTGCTGTTGCTTCCTTATCATGTTGAGGGATGTTAGGATATTG
TGGAACGGTGAAAGAACATCTGTTGCTCTGATGGTACTTGGAAAGTTGCTTGTCTTCT
GTGAGAACGGCTGTGGCCTTGGACATATGAACAGGAACCTATGGTCCAGTACAATGGTCA
GATATGGTCACTTGAAGCCAGGATGACAATTATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTTGAATTTCAGAGAGAGTTAAGTGTGCACTATATTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCTGCTGTTAGAATTCCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGCACCTTATCAAGAGGGTTGGGAAGAAAATGATTCCCT
TTTGAGAGGAACCAACAACTGCAGGTGCTGAGGTTCTGGAATCTCATTGGGTGACACAAA
TCCTGGCCATGTTCTCACCATTACTCTGCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
ACAGACAAATGATGTCCTGAGAACATCTCAGCACCTGTCTGAGTCCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAACATCCTGAACACATCCATGGCAAACAGCTTAATACACACT
TTGAGATGGAGGAGTTAAAAAGAACATGTCAGAACAAAACCACAAACTTGTTTTATTGGACT
TGTGAATTGGAGTACATACTATGTGTTGAGAAATATGAGAAATAAAATGTTGCCATAAAA
TAACACCTAACGATATACTATTCTATGCTTAAATGAGGATGGAAAAGTTCTGATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTTAGCC
TGTGTATGACTTTACTGAACACAGTTATGTTTGAGGCAGCATGGTTGATTAGCATTCCGCA
TCCATGCAAACGAGTCACATATGGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
CTAGTATATAAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATT
CTCAGCGATCTATTCTCTGATGCTAAATAAAATTATATCAGAAAACCTTCAATATTGGTGA
ACCTAAATGTTGATTTGCTGGTTACTAAAATATTCTACCACTTAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGATTTTGTATATAAGCTGTGTTAAATCTGATATAATTCACTG
TTCAGTTCTGATAATGTTAGAATAAACATTATGAAAAGAAAATTGTCCTGTATAGCATCATT
ATTTTTAGCCTTCTGTTAAATAAGCTTACTATTCTGTCCTGGCTTATATTACACATATAAC
TGTTATTAAATACTTAACCACTAATTGAAAATTACCACTGTTGATACATAGAACATTATTC
AGAATGTTGAGTCTGGCTTTAGGAAGTTAATAAGAAAATTGACACATAACTTGTGATT
AAGGACTGTGCTGTTCTCCAAATGAAGACTTTTGACACTAAACACTTTTAAAAA
GCTTATCTGCTCTCCAAACAAGAACAGCAATAGCTCCAGTCATATAAAATTCTACAGAAAA
TAGTGTCTTTACTCCAGAAAATGCTGTGAGAACATTTAACATGTGACAATTAGAGATT
CTTGTGTTATTACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTACAC
GAGTATAGTATATTATTGAAATGGGAAAAGTGCATTACTGTTGATTTGTGATTGTTAT
TTCTCAGAATATGGAAAGAAAATTAAATGTGTCATAAAATTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTATRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMWDPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPDTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGGACACCTTATCCCACTCATCCTCATCCTCTGATAAAGCCCCTACCAGTGCT
GATAAAAGTCTTCCTCGTAGAGGCCTAGAGGCCCTAAAAAAAAGTGTGAAAGAGAAGGGGACAAGGAACA
CCAGTATAAGAGGATTTCAGTGTTCCTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
CCTCTCATCACAGGCCACCTCGTGTACCCGTGCCCTAGATCCTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAAGTGGATGAGTCAGTCAGTCCTCTATGTGACAACCATGTGAATGGGAG
TGGTACCACTTCACGGCATGGCGAGATGCCATGCCATACCTCTGCATACAGAAAACCAGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCAGGGATTGTGCAACGCCAGGCTGTGCCAGCT
TCAATGGAACTGCTGCTCTGGAACACCACGGTGGAAAGTCAAGGCTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGCTGCTTCCACGTACTGTGGTCAATTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATAACCAGCGAGTCACATGCGCTCCAGGAACGTGTGCTAGGCCCTGACAGGCAGACATGCTTG
ATGAAAATGAATGTGAGCAAACACGGTGGCTGAGTGTGAGATCTGTGTAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCGTGTGCTAAGAAGTGTGAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTGGCTTGGATCTGAGAAAGGCTACAGTGTGAATGTCCCCGGGCCTGGTGT
CTGAGGATAACCACACTTGCACAGTCCCTGTGTGCAAAATGCCATTGAAAGTGAACATCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTTCTGACCAACACTCCCTGCCAGGAGTGTCAACGCCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTGTGATGACAAGATTGTGGCAGCAACCTCGTGA
CAGGTCTACCAAAAGCAGACCCGGGGAGCAGCGGGACTTCATCATCCGAACCAAGCAAGTGTGATCCGGT
ACCTGCGAGTTCCACGCCTGTACACCATTCTGAAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCGAAATCATGGGATCTTCCCAATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCAAGCTCGTACTCCCTACTTGGCATTGAGCCGTGGTGCACGTGAGCGGC
TTGGAAGCTTGGTGGAGAGCTGCTTGCACCCCACTCAAGGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAAGATGACTCGGTAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCAGG
TCCCTGTCTCAAGTTGTGGCAAGACCACAAGGAAGTGTCTGCACTGCCGGGTTCTGTCTGTGGAGTG
TTGGACAGCGTTCCCGTGTGCCACCGCGAATGCCGTGCGGGCAGGAGGAGAGACTCAGC
CGGTCTACAGGCCAGCGCTAACAGCGGCCGATCCGCACTGGAGACTAGTGCTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTTGGAGCTCTCCCCCACCAGGCTTAAGAACATCTGCCAACAGC
TGGGTTCAAGACTTCAACTGTGAGTTCAAGACTCCAGCACCACACTCTGATTCTGGTCCATTCACTGGCA
CAGGTCAACAGCACTGCTGAACAATGTGGCTGGGTGGGTTCATCTTCTAGGGTTGAAAACAAACTGTCCA
CCCAGAAAGACACTCACCCATTCCCTCATTTCTTCCCTACACTAAATACCTCGTGTATGGTGCACATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAACCCAGAAAAAAATAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCTTAAATATGTAATTAGTTACCTGAAATTCAATTCAAATGCAGACTAA
TTATAGGAATTGGAAGTGTATCAATAAAACAGTATATAATT

FIGURE 110

MPPFLLTCLFITGTSVSPVALDPCSAIISLNEPWRNTDHQLDSEQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCLWNNTVEVKACPGGYYVYRLTKPSVCFHV
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGCASHCLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFLTNTSCRGVSNGTHVNILFSLKTCGTVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFITLEIFKDNEFEEPYREALPTLKLRDSDLYFGIEPVVHV
SGLESILVESCFATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQQQLTGGSPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGTCAAGCGGACAAGGATGCTGGGCGTAGGGACCAAGGCCTGCCCTGCACACTGG
GCCTCCCTCAGCCAGTGCTGACCAGGGACTTCCTGACCTGCTGGCAGCCAGGACCTGTGAGGGAGCCCT
CCTGCTGCCCTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGTGGGATCCCCATCATCATAGCACTACTGAGCCTGGGAGTATCATCATATGT
GGTTGCTCTCATCAAGGTGATTCTGGATAAATACTACTTCCTCTGCAGGCTCTCCACTTCATCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCTTGGGGAGGACGAGGAGCACTGTGTCAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCGCCTCTCAAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACTGGTTCTCTGCCTGTTGACAACCTTCACAGAACGCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGATCGGAAACTCAAGTGGCCCTGTCCTCAGGCTCCCTGGTCTCCCTGCACACTGCTTGCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGGAGGAGGACATCCTGGACCCCCACTGGTCTCAGGGCA
TCAGCATCCAGTACGACAAACAGCACGTCTGGAGGGAGCATTGGACCCCCACTGGTCTCAGGGCA
GCCCACTGCTTCAGGAAACATCCGATGTGTTCAACTGGAAGGTGCCGGCAGGCTCAGACAAACTGGCAG
CTTCCCATCCCTGGCTGTGGCAAGATCATCATATTGAATTCAACCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTCCCCTCACTTCTCAGGCACAGTCAGGCCATCTGCTGCCCTTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATATTGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGTCGACATACTGCTGCAGGCGTCAGTCAGGTCATTGACAGCACACGGTGCAATGCAAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGTCAGGCATCCGGAAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTACCAATCTGACCAAGTGGCATGTGGTGGCATCGTTAGCTGGGCTATGG
CTGGGGGGGCCGAGCACCCAGGAGTATACACCAAGGTCTAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTTAATGCTGCTGCCCTTGCAGTGTGCTGGAGCCGCTCCTGCCCTGCCACCT
GGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTGGTACACCCCTGCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGCCTCAATTCTGTAAGAGACCCCTCGCAGGCCAGAGGCGCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTCTGCCAGCATCCAGGGAGAGACACAGCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTGTAAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTCTGCCAGGCCAGGGAGAGACACAGCCACTGAACAAGGTCT
GCCTACTAGAGCAAGAACCAAGTGTAAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTATTGTTATTACAGCTATGCCACTATTATAAGAGCTGTGTAACATCTGGCAAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIALLSLASIIIVVVLIKVILDKYFLCG
QPLHFI PRKQLCDGELDCPLGEDEEHCVKS FPEGPAVAVRLSKDRSTLQVLD SATGNWFSACFDN
FTEALAE TACRQM GYSRAVEIGPDQ DLDV VEITENS QELRMRNSSGPCLSGSLVSLHCLACGKSL
KTPR VVGEE ASVDSWPWQVS IQYDKQHVC GGSILD PHWVL TAAHC FRKHTD VFNW KV RAGS DKL
GSFPSL AVAK III IFNP MYPKDNDIALM KLFPLTFSGT VRPICL PFFDEELTPATPLWIIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT CQGD SGGPLMYQS
DQWHVV GIVSWG YGC GG PSTPGV YT KVSAYLN WIYNV WKAEL

Transmembrane domain:
amino acids 32-53 (type II)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCAAGTGATCCACCCGCCCTCAGCCTCCAAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGCCCTGAACAACCTTTCACTGCTGCTACACATTAAAATC
TGAAC TGCTAGGATTCTGACTTATGCAGGTTCTAGTGCCTACTCCTACCTACATTAAAATC
TGTTTTGTTCTCTTGTAACTAGCCTTACCTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCCTCGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCTCGCTGGAGCAGTGCCCTACCAACTGTCTACGTCT
GGAGGCAGTGACTCGGGCAGTGCAAGGTAGCTGAGCCTCTGGTAGCTGCGGCTTCAAGGTGGC
CTTGCCCTGGCCGTAGAAGGGATTTGACAAGCCCGAAGATTCATAGGCATGGCTCCACTGCC
AGGCATAGCCTGCTGTAGTCATCACTGCCCTGGGCCAGGACGGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACCGCTGCCGCCATCTAACCTTCACTGCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGCTTCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAAGTGTGCAAATTATGGGTAGAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGAGCTGTGGGAAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGAT
CCACGTACAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACATTCTCTCCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCCTACTAAAGCCACAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTCAAGCAAGGCCGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAAATGCCCTCAGTAAGCACAGGCTGCAAAATCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCAAGGGCCCTTATACCCAGGAGACTTGTGATTTGAAATTGAAACCCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTGGCTGTAATGCCAACATGG
TTTGGGAGGCCGAGGCCGGTAGATCACCTGAGGTCAAGGAGCTGAGGCTGAGGAAACGCCAACATGG
TGAAACCCCTGTCTACTAAAAAATACAAAAAAACTAGCCAGGCAAGGTGTGGCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTGAAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAACATTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA
TGGTTATTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLAALTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

© 2007 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.

FIGURE 115

CAGCAGTGGCTCTCAGTCCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTACACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAAATTGTCC
GTTTGCCCCAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGAAATGGCACTGATGAAACATTGAAAGTGCACGACTTTAAAACGGATAACACTGG
CATCTACTCGTGGGTCTCAAAAATGTTTATCAAAACTAGATTAAAGTGATTCTGAATT
CTGAACCAGAAGAGGAAATAGATGAGAATGAGAAATTACCAACTTCTTGAAACAGTCAGTG
ATTGGGTCCCAGCAGAAAGCCTATTGAAACCGAGATTTCTTAAATTCCAAAATTCTGGA
GATTGAGTAAACGTGACCATGTATTGGATCAATCCCCTCTAAATATCAGTTCTGAGTTACAAG
ACTTTGAGGAGGAGGGAGAAGATCTTCACTTCCGCAACGAAAAAAAGGGATTGAACAAAAT
GAACAGTGGTGGTCCCTCAAGTAGAAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTCCAATAATGACTATACTGAAATGAAATTGATCCATGCTGGATGAGAGAG
GTTATTGTTGATTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTGTGAACCTTTACTA
GGCTACTACCCATATCCACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCCTG
TAACTGGTGGTGGCCCGCATGCTGGGGAGGGTCTAATAGGAGGTTGAGCTCAAATGCTAAAC
TGCTGGCAACATATAATAATGATGCTATTCAATGAATTCTGCCTATGAGGCATCTGGCCCT
GGTAGGCCAGCTCTCCAGAATTACTGTAGGTAATTCTCTCTCATGTTCTAATAAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKSKICKSLKICGLVFGILALTLLIVLFWGSKHFWPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEEIDENEETTTFFEQSVIWWPAEKPIENRDFLKNSKILEICDNVTMYWINPTLISVSE
LQDFEEEEDLHFPAKEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMED
ERGYCCYICRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTCACACCTCGCAGCAGGGGGCAGCTTCGCAGGCCA
GGCGGGCGGCCAGGATCATGTCCACCACATGCCAAGTGGTGGCGTCCCTCTGCATCCCTGGGCT
GGCGGCTGCATCGCGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCCTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAGTTCAAGGCTTCAACGAATGCAGGCC
TATTCACCACATCCTGGACTTCCAGCCATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGCCT
GGGTGCCATTGGCCTCTGGTATCCATCTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGATCATGTTCAATTGTCTCAGGTCTTGTCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAATTGGATGTCCACAGCTAACATGTACACCGGATGGTGG
GATGGTGCAGACTGTTCAAGACCAGGTACACATTGGTGCCTGTTCTGGGCTGGTCGCTGGAGGCC
TCACACTAACATTGGGGGTGTGATGATGTCATGCCCTGCCGGGCTGGCACAGAGAAACCAACTACAAA
GCCGTTCTTATCATGCCTCAGGCCACAGTGTGCTACAAGCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAAAACAAGAAGATAACGATGGAGGTGCCGCACAGAGGACGAGGTACAATTTATC
CTTCCAAGCAGCACTATGTGTAATGCTCTAACAGACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCATCTAGATTCTTGCTTTGACTCACAGCTGGAAGTTAGAAAAGCT
CGATTTCATCTTGGAGAGGCCAATGGTCTTAGCCTCAGTCTGTCTAAATATTCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTAGCTCACATTTCATCCTCTATTCTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGTTTAATCTCTCTCACATTGATGATTTAGACAGACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCCAGCTATCCCCAGAAAATTGGAAAGGAAA
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGTCTCCCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATAATTGTAATCTCTCCAGCCCATGATCTGGTTCTT
ACACTGTGATCTAAAGTTACCAAACCAAAAGTCATTTCAGTTGAGGTTCTGAGCTCTCCACTGGAGTCCTTTCTG
TTGACATCTCTTATTACAGCAACACCATTCTAGGAGTTCTGAGCTCTCCACTGGAGTCCTTTCTG
CGCGGGTCAGAAAATTGTCCTAGATGAATGAGAAAATTATTTTAAATTAAAGTCCTAAATATAAGTTAA
ATAAAATAATGTTAGTAAATGATACTATCTCTGTGAAATAGCCTACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAGTCATGCTTAAGTACAATTCC
ATGAAAAGCTCACACCTGTAATCTGACTTTGGGAGGCTGAGGAGGAAGGATCACTGAGCCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAATCAGCCAGTCA
TGGTGGCATAACACCTGAGTCCCAGCATTCCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAATAATAATGAAACACAGCAAGTCCTAGGAAGTAGGTTAAAACATAATTCTTAA

FIGURE 118

MSTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMMGVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHHSVAYKPGGFKA
STGFGSNTKNKKIYDGGA
RTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTTCTGTGGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTCGGAGTCC
AGCTGGCTAAACATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGCTGTTCTTG
GTGGTGGTGAATGGTGGCACAGTGGCTGACTGTCAGCCTCAGTGGAGAGTGTGGCATCATT
GAAAACAACATCGTGGTTTGAAAACTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGGCCAGAG
GACTGATGATGCTGCTCCGTATGTCCTCTGGCTTCATGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTCAT
CATCACGGGCATGGTGGTGCATCCCTGTGAGCTGGGTGCAATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTGTTCTGCTGCGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTGATACCTCCCATCGCACAAACCCAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTATGTTTTAACTTACTATAAGCCATGCAAATG
ACAAAAAATCTATATTACTTTCTAAATGGACCCCCAAGAAACTTTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAACTGTGCATCAGCTATTGATTGATCTATAAGCTATTCAGCAGAATGAGATA
TTAAACCAATGTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATCTA
CTCTTTTATCATTACTTCAAATGACATTGCTAAAGACTGCATTTTTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATATGGT
TTTATTAAATGAAATGCCAGTCCATTACACTGAATAAAATAGAACTCAACTATTGCTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTAAATTGTTAAAAACAGCTTAGGGATTAATGCTCCA
TTTATAATGAAGATTAATGAAGGCTTAATCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTATCCTCTCTCCCAGAGGCTTTTTT
CTTGTGATTAAATTAACATTAAACATTAAACAGCAGATATTGTCAAGGGCTTGCAATTCAAACTGCTT
TTCCAGGGCTTACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGATGGTTAGGAAAGTG
AAAATTTTGTGTTTGATTTGAAGAAGAATGATGCATTGACAAGAAATCATATATGTATGGAT
ATATTTTAATAAGTATTGAGTACAGACTTGAGGTTCATCAATATAAAATAAAAGAGCAGAAAATA
TGTCTGGTTTCATTGCTACCAAAAAACACAACAAAAAAAGTTGTCTTGTGAGAAACTTCACCT
GCTCCTATGTGGTACCTGAGTCAAATTGTCATTGTTCTGTGAAAATAATTCCTTCTGTA
CCATTCTGTTAGTTACTAAATCTGTAATACTGTATTGTTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTGTCGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTTATTAATAAAATTGTACATTGTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSIFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLIALSPDLQAARGILMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDYNSIVNVAQKRELGEALYLWTTALVLIVGGALFCCVFCCNEKSSY
RYSIPSHRTTQKSYHTGKKSPSVRSQYY

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAGGCATTGATGCAGCCTGCCGCGCCCTCGGAGCGCGGGAG
CCAGACGCTGACCACGTTCCCTCCTCGGTCTCCCTCCGCCTCCAGCTCCGCCTGCCCGAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCCTCCCGCAGCGGCTCCGCCTCGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCGTCGAGCGCCTCTGAGATCCCCAAGGGAAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGTTACAAGGCCAGCAGGAGTGC
TGGTCGAGACGGGAGCCCTGGGCAATGTTATTCCGGGTACACCTGGGATCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTGAGGAGTCCTGGACACCCAAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTTGGAAAATTGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGCTTAAGAGTTGTCAGTGGCTACTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTCACATTCAATGGAGCTGAATGTTCAAGGACCTCT
ATTGAAGCTATAATTATTGGACCAAGGAAGGCCCTGAAATGAATTCAACAATTAAATT
CACTTCTCTGTGGAAGGACTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCTGG
TTGGCACTGTTCAAGATTACCAAAAGGAGATGCTTCACTGGATGGAATTCAAGTTCAGTTCT
ATTATTGAAGAACTACCAAATAAATGCTTAATTTCATTGCTACCTCTTTTATTATGCC
TTGGAATGGTCACTAAATGACATTAAATAAGTTATGTATACTGAAATGAAAGCAAAG
CTAAATATGTTACAGACCAAAGTGTGATTCACACTGTTAAATCTAGCATTATTCA
CTTCAATCAAAAGGTTCAATATTTTTAGTTGTTAGAATACTTCTTCAGTCACATT
CTCTCAACCTATAATTGGAATATTGTTGTGGCTTTGTTCTTAGTATAGCATT
AAAAAAATAAAAAGCTACCAATCTTGTACAATTGTAAGAATTGTTAAGAATT
TAAATAAAAATTATTCCAACA

FIGURE 122

MRPQGPAASPQRRLGLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFGSGLRLKCRNACCQRWYFTFNGAECGPLPIEAIYLDQGSPEMNSTINIHR
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGGCGGTACGGGGCTCTCCTGCCTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGAACTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATAACATCATGTTTCGATAAGAAGAAATTGAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCAAAAAAACTGTAAGATGCAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTGATGTTGCTGCGAATGCCGTGTTGGGATTATTGTTCTGGAG
TGTCTGCGTGGCTGGCAAAGAATAATGTCCAAAATCGTCCATCTCCAAGGGTCCAATT
TCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTCAGCTGACAGGGCTGTCATGCACTG
GCCCTAAGCCAAGCAAAGACCTAACGGACGACCTTGACAATACAAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTATAGCCCCACTGTCTTACTGACAATG
CTTTCTCTGCCGAACGAGGATGCCCTAACGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTA
ATCTCAGAAATTACAGGAGATACCCCTCAAGTATATGCTGGTTGCTTAGGTTGTCCTCGCT
ATAACAGCCTTAAAAACTTAAGTATAATCAATTAAAGGGCTAACCGACTCACCTGGCTATAC
CTTGACCATAACCATACTAGCAATTGACGAAATGCTTTAATGGAATACGCAGACTCAAAGA
GCTGATTCTTAGTCCAATAGAATCTCCTATTCTTAACAATACCTCAGACCTGTGACAAATT
TACGGAACCTGGATCTGCTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTGGGGCTTG
CGGAAGCTGCTGAGTTTACATTACGGCTAACTCCCTGAGAACCATCCCTGCGAATATTCCA
AGACTGCCGCAACCTGAACTTTGGACCTGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGCTGGCATGACTCAAAGAACCTCACCTGGAGCACATCAATTCCAAGCTCAAC
CTGGCCCTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTACTTGCACTGGAATAAAATCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTTACAAGGCTGATTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCCTGAATCTGCAGCGCTCAACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATTTGGATTCCTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATGGGAATGCAAGCAGAAATATTGCTCCCTGTTAAACTGGCTGAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCACTGAAAGACTACAGCATCTGTGGCAAAGTACTACAGAGAGGTTGATCTGGCAG
GGCTCTCCCAAAGCCGACGTTAAGCCCAAGCTCCCAGGCCGAAGCATGAGAGCAAACCCCTT
TGCCCCCGACGGTGGGAGGCCACAGAGCCCCCAGAGACCGATGCTGACGCCAGCACATCTCT
TTCCATAAAATCATCGGGCGAGCGTGGCCTTTCCGTGTCGTGTCATCCCTGCTGTTAT
CTACGTGTCATGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAAGCGCTCCCTCATCGAA
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTATGTA
GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAAGAGCTTAAAGCT
GGGAAATAAGTGTGTTTATTGAACTTGGTGAATCAAGGGAACGCGATGCCCTTCCCT
TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTGTCCGTTTAGTGCATTCAATAACT
GGTCATTTCTCTCATACATAATCAACCCATTGAAATTAAATACCACAAATCAATGTGAAGCTT
GAACCTCCGGTTAATATAATACCTATTGTATAAGACCCCTTACTGATTCCATTATGTCGCA
GTTTAAGATAAAACTCTTCATAGGTAAAAAAAAAA

FIGURE 124

MGFN VIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGMVYCESQKLQEIPSSISAGCLG
LSLR YNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHRSNSLRTIPVRI FQDCRNLELLDLGYNRIRS
LARNV FAGMIRLKELHLEHNQFSKLNLA LFPRLVSLQONLYLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLTFIGQEILD SWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELOGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKKIAGSVALFLSVL VILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKKRQSLKQMTPSTQE FYVDYKPTNTETSEMLLNGTG PCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCCGTCCCCGAGGAGGAGGAGGGCTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCCGGCTGCCGGCTACCGTGGCCGAGCT
AGCAACCTTCCCTGGATCTCACAAAACCGACTCCAATGCAAGGAGAACGAGCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCTATAGGGATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCATTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTATTGGCCAGTT
TTAGCCAATCCAACGTACCTAGTGAAGGTCAGATGCAAATGAAAGGAAAAGGAAACTGGAAGG
AAAACCATTGCGATTCGTGGTACATCATGCATTGCAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGCTGGTACCCAAATACAAAGAGCAGCAGTGGTAATATGGGAGATTAA
ACCACTTATGATAACAGTGAACACTACTTGGTATTGAATACACCACCTGAGGACAATATCATGAC
TCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTCTATTCTGGAACACCAGCCATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAGGAAGGGACTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGTGAAGGATTGAGTCTATATAAAGGCTTTTAC
ATCTTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLCVVFGLKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRI MNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSLRLMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTGGCGGCCGGCAGGAGAGCGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTGGGCCTAGTCCAGCGAGGCGGACGGCGGCCGTGGGCCCATGGCCAGGCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGTGGTGACGGGGCCTGGGGCATCGCGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTCAGGCTACCCGGACTTTGATCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTGGCCCGGCCTGACACCCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCATGA
AGGAGCGGAATGGGACATGGGACATCAATTAAACATCAATAGCATGTCGGCCACCGAGTGTAA
CCCCTGCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTCACTGCCGTACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCGAGGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCCTCAAACCTCACGACAAGGACCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGAGCTCC
TCCTTCCCTCCCCACCCCTCATGGCTTGCCCTGCCCTGGATTAGGTGTTGATTCTGGAT
CACGGGATAACCACTCCTGTCCACACCCCGACCAGGGCTAGAAAATTGTTGAGATTTTATA
TCATCTGTCAAATTGCTTCAGTTGAAATGTGAAAATGGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTTGCCCTGGCACTTGGCCTTGCTGCTCTCAGTG
TCTTCCCTTGACATGGAAAGGAGTTGTGGCCAAATCCCCATCTTGTGCACCTCAACGTCTG
TGGCTAGGGCTGGGTGGCAGAGGGAGGCCACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCCTCTGCCCTGCCACTGCACCCCTCTCCCTTATCTATCTCCTCTCGGCTCCCC
AGCCCCAGTCTGGCTTCTGTCCCTGGGTCACTCCCTCACTCTGACTCTGACTATGGCAG
CAGAACACCAAGGGCCTGGCCAGTGGATTTCATGGTGATCATTAAAAAGAAAAATCGAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDSLNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLPLSVTHFYSATKYAVTALTEGLRQELREAQTHIRATC
ISPQGVVETQFAFKLHDKDPEKAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGGTCTTCCTCAGCCTCCTGCCGGTGGCTACACCAT
CATGTCCTCCCACCCCTCTTGACTGCGGCCGTTCAAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCTCCGAGGCAGTCGCTCAGAGGGCCTGCCAGAATTCCAGTTCTGGTTCATGC
CAGCCTGTAAAAGGCCATGGAACTTGGGTGAATCACCGATGCCATTAAAGAGGGTTCTGCCA
GGATGAAATGTTAGGTCGTTCTGTCTGCGCTGTTCAATTCAAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAAATTCTCATGTATTTCATTTCAATTATTAA
TTAATTAACTGATAGTTGTACATATTGGGGTACATGTGATATTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCAGGTTCAAGCGATTCTCATGCCCTCACCTCCAAAGTAGCTGGACTACAGGCAT
GCACCAAAATGCCCAACTAATTGGTATTTAGTAGAGACGGGTTTGCCATGTTGCCAGG
CTGGCCTGAACTCCTGGCCTCAAACAATCCACTTGCCCTGGCCTCCAAAGTAGTTATGATTACA
GGCGTGAGCCACCGTGCCTGGCTAAACATTTATCTTCTTGTTGGAACTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGACTTCTCCCTCT
ATCTAACTGTATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCACCTCCATGAGATCCACTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGCTTTCTGTGCCTGGCTTATTCACITTAACATAATGACTTCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTCGTTCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLVLFLSLLPVAYTMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGHGTLGESPMPFKRVFCQDGTVRSFCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCTCATCGGGACTAAAATTGGGCTTCCCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGTCGCTGCGATGCCGGTTTCAATTACTGTAATGATCGCTTCTGACATCCATTCAAACAG
GAATACCAAGAGGATGCTACAACCTCTACCTCAGAACACAAATAATAATGCTGGGATTCCCTCAGAT
TTGAAAAACTTGCTGAAAGTAGAAAGAATATACTTACACAAACAGTTAGATGAATTCCCTACCAACCT
CCCAAAGTATGTAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGGAAGAATTACATTTAGATGACAACCTGTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTCCCTGTCCCGTAATCACCTTAGCACAATTCCCTGGGGTT
GCCAGGACTATAGAAGAACTACCGCTGGATGATAATCGCATATCCACTATTTCATCACCATCTTCAAG
GTCTCACTAGTCTAAAACGCCTGGTCTAGATGAAACCTGTTGAACAATCATGGTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAATTGACAGAGCTGTCCCTGGTGCAGGAAATTCCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTTCAAGATAACCACATCAATGGGTGCCCTCAAATGCTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCATAATAACCTAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACGTTCTCGCAACATCCCTGGTATTGCGGGTGCAGATGAAATG
GGTACGTGACTGGTACAACTACACTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCGTGGGATGGCTATTAAAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCACACTGCAATACCCAACACAGTGTATCTGCCAAGGACAGTGGCAGCTCCAGTGAC
CAAACAGCCAGATATTAAAGAACCCCAAGCTCAACTAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAAA
CAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTCAATCTCTGGAAACTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTAAACTGGGCCATAGCCCGATTGGATCTATAACAGAAACATTGT
AACAGGGGAACGCACTGAGTACTGGTCACAGCCCTGGAGCCTGATTCCCTATAAAAGTATGCATGGTC
CCATGGAAACCCAGCAACCTCTACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCTACAACCACCCCTCAATCGAGAGCAAGAGAAAGAACCTACAAAACCCCAATTAC
TTGGCTGCCATATTGGGGCTGTGGCCCTGGTTACCATGGCCCTTCTGCTTAGTGTGGTATG
TTCATAGGAATGGATCGCTTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAACGCTGGCACTAAGAAGGACAACCTATCCCTGGAATCAGGGAAACTCTTTCAAGATGTTACCAAT
AAGCAATGAACCCATCTGAAGGAGGAGTTGTAATACACACCATATTCCCTCTAATGGAATGAATCTGT
ACAAAACAAATCACAGTGAAGGAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGGTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPS DILKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDNSVS AVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNL PQGIFDDLDNITQLILRNNPWYCGCKMKWV RDWLQLSPV
KVNRG LMCQ APEKVRGMAIKDLNAELFDCKD SGIVSTI QITTAIPNTVYP AQQWPAPVTKQPD
IKNP KLT KDQQT GSPSRKTITITVKS VTS DTIHISWKLA LPMTA RL SWLK LGH SPAFG SITET
IVTGERSE YLVTALEPDSPYKCMVPMETSNLYLFDETPVC IETETAPLRMYNPTT LNREQEKE
PYKNPNLPLAAIIGGAVALVTI ALLALVCWYVHRNGSLFSRN CAYSKGRRRKDDYAEAGTKKD NS
ILEIRETSFQMLPISNEPISKEE FVIHTIFPPNGMNLYKNNHSESSSNRSYR DSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,

640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCCAGAGTCCTTGCCCAGGCCACCCAGGTTCTGGCA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATGCAAGACCTGATCCTGTCT
GCGCTGGAGAGAGGCCACCGTCTCCTAGAACAGAGGCTGCCGTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGTGAGAAGCAGCTAAAAAGTGTCCGGAGAAGTGGGCCAGGAGGCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGTGATCCCAAGTACCTAACAGAGAGTCCAGCTGACCCCTCCAGCCGG
GTTTTGGAAGCTCCCACATGCCCTGGATCCACACTGATGCCCTTGGTGTACCCCACGTTGGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGAACGTGTGCCCTGGTGCAGCTGCTGGAAACCGGGACG
GACAGCAGCGAGCCCTGCCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCCTGCCCACCAACTGCTCTTCCCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATGGATACGCCAACCTACCCGGACATCTTATGGAAAACATCATGTTCTG
TGGAAATGGCGGCTTCTCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGAGTGAAGAGGCGAGAAAACAATTCCAGATTCTGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTCACTGCAACCTTGCCTCTGGTTCAAGC
AATTCTCTGCCTCATCCTCCGAGTAGCTGGACTACAGGAGCGGCCACCATACCTGGCTAAT
TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGTCTCGAACCTCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCAAAGTGTGGATTTAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAAGGGCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCCCCTAGATGGCTGCCCTCCCACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATACCCCCCAGCAAACAGAGAGGCCACACCCATCCACACCG
CCACCAAGCAGCCGTGAGACGGACGGTCCATGCCAGCTGCCCTGGAGGAGGAACAGACCC
TTAGTCCTCATCCCTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAAATCTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAGTCAACTGCAACTGAAACTGAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLARAQGAERVTSSEQRPAMASLGLLLLLTLAPPLWSSSLPGD
TAESKATIADLILSALERATVFLEQRLPEINLDGMGVVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCGYCLSHQQLFFLWARMRGCTQGPLQQSQD
YINLFCAANMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGCCGCTCTGTGGGCTCTTCCCCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTGGAGGTGCCGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGTGCGCATAGGAGATCGCTTCAAGATTGAGGGCGTGCAGTTGTTCCAGGGTGAAGC
CTCAGGACTGGATCTGGCGGGCCGAGTGTAGACGGAGAAGAGCACGTGGTTCTTAAG
ACAGATGGGAGTTTGTGGTCATGATATACTTCTGGATCTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTGATCCCCTCGAGTGGATATCACTTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAACATCAGAGGTTGTCAACTGCCCTATCCTCTCCAAATGAAATCTCAGGT
CCACCTTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTTCTAATGAACCAATGGT
TATGATGATGGTCTTCTTATTGATATTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCAATGAATATGCTGAATTCAACCAGTGGCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
AAAAACAGGCAAAGTGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGGAAAACCGTGTGAAGCAACTACTATAACTT
GAGTCATCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTTGTACTTGGTACACGAGAAAACCCAGCTTCATCTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAAATTACAGTGTCTTATAGAAAATGCCATTAATAATTAT
ATGAACTACTATACTTATGTATATTAAATTAAAACATCTTAATCCAGAAATCAAAAAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEHVGFGLTDGSFVVHDIPSGSYVVEVVSPAYRFDPVRVDITSKGKMRARYVNYYIKTSE
VVRLPYPLQMKGSSGPPSYFIKRESWGTDFLMNPVMVMMVLPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCAGTTCTCCCCAGTCCCCGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCTGGAAGGGAGGCTATATCGTCAATTCCCCAAACAA
GTTTGACATTCCTCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGACTAACGGCAGGATGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTCTTCACGGGAGGCTTGGCAGT
TTTCCTTACTCCTGTGGTCTCCAGATTCAGGCCTAAGATGAAAGCCTCTAGTCTGCCTTCAGC
CTTCTCTGCTGCGTTTATCTCTATGGACTCCTTCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTCAGGAAATACGAAATGGATTCTGAGATACGGGCAGTG
TGCAAGCCAAGATGGAAACATTGACATCAGAACTTAAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAAGATCAGCAGCCTGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATAACGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTGGGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAAATAGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAGAGGAGGCATGCCCAA
ACCACCATCTCTTACTGTACTAGTCTTGCTGGTCACAGTGTATCTTATTGCATTACTTG
CTTCCTGCTGATGATTGTCTTATGCATCCCCATCTTAATTGAGACCATACTTGTATAAGATTT
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTATTATTGCTATTAA
ATGTATTTATTTTTACTTGACATGAAACTTTAAAAAAATTCAAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAAAAACCTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCATTGTATTCAACTAAGGACATATTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTGTGATGTGAAATTGCAC
ATCTACCTTACAATTACTGACCACCCAGTAGACTCCCCAGTCCCATATTGTGTATCTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTTCTTGCATACCAAAAAAAA
AAAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMASSLAFLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGTCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTCCGCATCCAGCC
TAGCGTGTCCACGATCGGGCTGGCTCCGGACTTCGCTACCTGTTGGTAGCGATCGAGGTGC
TAGGGATCGCGGCTTCCTCGGGATTCTCCCGCTCCCGTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCCAGGCCAACCCTGGCTGGAGCCAGTCTAACTGGACCACGCC
ACCACCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCCTGAGAGATGATTTGTTG
GGTCAAAGGGTGAAATTATGCCCTACACAACCTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTGAATCAAGGCATTGATGACGGGAG
CCPTCCCTGGTTTGTGACGTCACTCAGGAACCTCAATTCTCCTGCACTGCTGGAAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTATGGAGATGAAACCTGGGTTAATT
TTCACAAAGCATTGTGGAATATGATGGAACAACCTATTTCTGTCAGATTACACAGAGT
GGATAATAATGTACAGGAGGCAATTGGATAAGTATTAAAAGAGGAGATTGGGACATATTATCC
TCCACTACCTGGGCTGGACACATTGGCCACATTCAAGGGCCAACAGCCCCCTGATTGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTACTGCACTGCGAAGGAGAGA
GACGCCCTAACCAATTGCTGGTCTTGTGGTGAACATGGCATGTCAGAACAGGAAGTCACG
GGGCCTCCACCGAGGAGGTGAATACACCTCTGATTAAATCAGTTCTGCGTTGAAAGGAAA
CCCGGTGATATCGACATCCAAAGCACGCTCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAGACAGTGTAGGGAGCCTCTTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGAATACAGTGCAGCTAGTAAACTGTTGCAAGAG
AATGTGCCGTATGAAAAAGATCCTGGTTTGAGCAGTTAAATGTCAGAAAGATTGCATGG
GAACTGGATCAGACTGTACTTGGAGGAAAGCATTCAAGTCCTATTCAACCTGGCTCCAAGG
TTCTCAGGCGAGTACCTGGATGCTCTGAGACGCTGAGCTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCCCTGCTCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAAGTCCA
CTGTCATCTCTGGTTTCTGCTCTTATTGGTGAATCCTGTTCTTCCGGCGTTACGT
CATTGTGTGCACCTCAGCTGAAAGTCTGCTACTTCTGTGGCCTCTCGTGGCTGGCGCAGGCT
GCCTTCGTTACAGACTCTGGTGAACACCTGGTGTGAGCTGCCAGTGGCTGCCCTGGAC
AGGGGCTCAGGGAAAGGACGTGGAGCAGGCTTATCCCAGGCTCTGGGTGTCCGACACAGGTG
TTCACATCTGTGCTGTCAGATGCCCTAGTTCTGGAAAGCTAGGTTCTGCACTGTTAC
CAAGGTATTGTAAGAGCTGGCGTCACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCAGCAGAGGTGTGGAGCTGAGCTGAGGGAAAGAGACAATCGGCCCTGGA
CACTCAGGAGGGTAAAGAGACTTGGTCGCACTCATCCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTCTCAAGGGACGTTCTGTTGGAATTCTAGTCCCTGGCC
TCGGACACCTCATTGTTAGCTGGGAGTGGTGGTAGGCAGTGAAGAAGAGCGGATGGTCAC
ACTCAGATCCACAGAGCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGCCCC
ACCCCAACCCCTGCACAGCCCTCATCCCCTTGGCTTGAGCCGTAGAGGCCCTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTATCAGGGCACAGGCTCTCGGAGCCAGGATGATCTGT
CCACGCTTGCACCTCGGGCCATCTGGGCTCATGCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACAAAAGAATAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSQTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPLF
SKVIVILIDALRDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWKLFPKHFVEYDGTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGTTATCGTGACGCACCTGAAAGCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGGAGTACAACCT
TCCCTATAGAAAACAAC TGCCAGCACCTTAAGACCACTCACACCTCAGAGTGAAGAACCTAAAC
CCGAAGAAATT CAGCATT CATGACCAGGATCACAAAGTACTGGTCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAA ACTACATACGCCAGAGATCTTCTTGCAATTAGCCTCATCCTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGGAGTTTGCTCTAC
TGTGACAAGGATAAAGGACAAGTCATCCATCCCTCAGCTGAAGAAGGAGAAACTGATGAAGCT
GGCTGCCAAAAGGAATCAGCACGCCGCCCTTCATCTTATAGGGCTCAGGTGGCTCCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCGGATGGTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTACAGATAAATTGAGAACAGGAAACACATTGAATTTCATTCACCAGTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCA CGCATTTAGGAAACTGCCCAT TGAACGCCCTCGCTA
ATTTGAAC TAATTGTATAAAAACACCAAA CCTGCTCACT

FIGURE 142

MLLLLEYNFPIENNQHLKTTHTFRVKNLNPKKFSIHDDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSNMLESAAHPGWFIGTSCNCNEPVGVTDFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCCTCCCTTAA
TCCAGGATCCTGTCCTCCTGCTGTAGGAGTGCCTGTCAGTGTGGGGTGGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGTGCCTACAGTCAGCACCATGCTGGCCTGCCGTGG
AAGGGAGGTCTGTCCTGGCGCTGCTGCTGCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTTCACCGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCAACAGACTGGGG
CACATCTGAATTCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAACGCCATTCCAAGAAAGCACAG
AGCTGAACAATACCTCACCTGCTTCTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGAGGGATTCCACTGAGTGAACCCACTCACAGGCTTGCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTAGAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGGCCAAGGTTCTGACCGGATGAGGAAG
CACCTGAGCTGGTGGCTGCCACTGTCATGCTGCTTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTCATCAAGCAAGGCCGCAAGCTGACATT
GACTTCGGAGGCCAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCACCGCTGCATCAATGCCA
CCCAGGCCGAACCAGGGGAGTCCAGAACGCCAGAACAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTTGCTCCCTCAAGCATTGCGAGTTTGGAGAGGGCGCAGGACT
TCGGGTCAACCATGCACCAGCCAGTGCCTCTGCCTTCTGGTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCAACCCACGCGTGTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCCGCTCCACGTATGCGCCCTGGTATGCGCTGCGTTCTGATAAGATGGGGACTGTGGCTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTCACAGTACTTCCAACAACTCTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGC
ACCAGCTAGAAAATGGCAGAGCCAGGATTCAACCCTGGCTTGTCTAACCCAGGTTCTGCTCT
GTCCAATTCCAGAGCTGCTGGTGTCACTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNLHQQ
VLWRLVQELCSLKHCEFWRERGAGLRTMHQPVLLCLALIWLVMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCACCATGTCAGGTCCAGAGTCTCATTT
CCTGATGATTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTTCTCTGCCTCCTCT
CTGTCTCTTCCCCTTTCTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCACTGGGTAGGTACATGTGAAGGTTGTTATGAGTATATTGCCATGATGGCTGAGG
CAACTTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCCATGATGGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCC CGCGCCCCAACCCCTGCTTATCCCTGACCGTCAGTGTCAAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCCGCCCCAACACCACCCACCCCTCTGGCTCTTCTGTTTACTCCTCTTTCTTCA
ACAAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGCTGTGACCCAAGCCGAGCGTGGAGAATGGGGTT
CCTCGGGACCGGCACCTGGATTCTGGTGTAGTGTCCCATTCAAGCTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAAGTGCAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAACATATCCTCAGAAAACAAGCCAGGTAGAGCAACTATTCTT
TGTTGATAACTTGAACCTGCTAAAGGCATAAACAGAAAAGGAAAAATTGAGAAAGAAAAGACAATCTA
TAAGAACGCTCCCCACTTGATAATAAGTTGAATGTGGAAAGATGTTGATTCAACCAAGAACATCGAAAAGT
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTAACCGCTGAAGACATTGTCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTGGCCTTATCACAGAAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAATTAACTCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAATACCAGAGAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTGCTAAGGGAGAAAAGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAAGGAGAACTAAACCTACAGTGAAGACAACATTGAGGAACCTCA
ATATTTCCAAATTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAGCAAAGAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTGTGAAGATGATGGTGAATATGGAACAAATATCT
CCAGAAGAGGTGTTCTACCTTGAAACTGGATGAAATGATTGCTCTTCAGACCAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCAGCACCCTAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAACAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAACGCTATTGGAAGC
CATCAGAAAAATATTGAATGGTGAAGAAAACATGACAAAAGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAAAACAGCTGATGCTTATGTGGAGAAAGGCATCCTGACAAGGAA
GAAGCCGAGGCCATCAAGCGATTATAGCAGCCTGAAAATGGCAAAGATCCAGGAGTCTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTCTAATTCTGTGATTAAAATTTTGACCCAAGG
GTTATTAGAAAGTGTGAATTACAGTAGTTAACCTTACAAGTGGTTAAAACATAGCTTCTTCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAA
AAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKYPHENPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEVAE
VLQKLISKEANNYEDPNKPTSWTENQAGKIPEKVTMMAIQCGLAKGENETVSNTLTLNGLE
RRTKTYSEDNFEEQLQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFPAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKRNEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGAACATTCTGAGGGAGTCTACACCCGTGGAGCTCAA
GATGGTCCTGAGTGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCAATTAAAGGTGAAGAGATCAGC
GTGGTCCCAATCGGTGGCTGGATGCCAGCCTGCCCCGTATCCTGGGTGTCAGGGTGGAAAG
CCAGTGCTGTATGTGGGGGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCAAGGAATCCAAGAGCTCACCTCACCTCACCGGCGGACATGGGCTCACC
TCCAGCTCGAGTCGGCTGCCAACCCGGCTGGTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCACTCACCAGCTTCCCAGAATGGTGGCTGAAATGCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACCTCCCTGGCAGAGCCAGCTGGGTGAGGGT
GAGTGGAGGAGACCCATGGCGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTGTCTCTGGTCCCAGTTGGATAAATTCTGAGATTGGAGCTCAGT
CCACGGTCCTCCCCACTGGATGGTCTACTGCTGTGAAACCTTGTAAAAACCATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGGTGGGGGAGTGGTGGGAATCATTCCTGCT
TAATGGTAACTGACAAGTGTACCCCTGAGCCCCGCAGGCCAACCATCCCCAGTTGAGCCTATA
GGGTCACTAGCTCACCAGTGAAGTCCTGCACTCACCAGTGTGCAAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTGGCCAGCCCCACCCCTTCCCTTAATCCTGCCACTGTCATA
TGCTACCTTCCTATCTTCCCTCATCATCTTGTGGCATGAGGAGGTGGTGTGAGAA
GAAATGGCTGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTAAAAACCAA
GATAACAATCAAATCCCAGATGCTGGTCTTATTCCATGAAAAGTGTCTGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTATTAAATTAAAGATAACCTATTATATT
TCTTTATAGAAAAAAAGTCTGAAAGAGTTACTTCAATTGTAGCAATGTCAGGGTGGCAGTAT
AGGTGATTTCTTTAATTCTGTTAATTATCTGTATTTCTAATTTTCTACAATGAAGATGA
ATTCCCTGTATAAAAAGAAAAGAAATTATCTGAGGTAAGCAGAGCAGACATCATCTGA
TTGTCCTCAGCCTTCACTTCCCAGAGTAAATTCAAATTGAATCAGCTCTGCTCTGGTTGG
TTGTTAGTAGTGTAGTTAGTTAAGACAAGGTATGCTGGATGAAGGTAGACCTAAATTCAATAT
GTGGCTGGAATCTGGTAAGGAACCTTAAAGAACAAAATCATCTGTAATTCTCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCAGCTCTAAGAACAGGCTGTACTGGTTGA
ATTGTGTCCTTCAAAATTCAATCCTTCTGGAATCTCAGTGTGAGTTATTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGCAGAGATGGAGTTTGCAGCCACAAGCTAAGAAAACCCAAGGATTGGCAACC
ATCAGAACGTTGAAAGAGGCAAAGAAGAAATTCTCCCTAGAGGCTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAAACGAAGAACATAAAATTCCGGCTGTTAA
GCCACCAAGATAATTGGTTACAGCAGCTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
CTCCTCGTGTACATTCTGTTGTCAGCCACAAGCTAAGAAAAGTGTCTTGTGACCAA
TAGAATATGCCAGAGTGTGATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC
TACTTGACCCCTCTCTCTGCCACCCACCGCCCCAATCTATCTGGCTCACTCGCTCTGGGG
AAGCTAGCTGCATGCTATGAGCAGGCCTATAAAGAGACTTACGGTGGAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATGATGTTGTTTT
AAGTTGCTCAGTTGGTCTAATTGTTATGCAGCAATAGATAAAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTCCTTATGGGGACCTGGCCACCAGCTGCCTCCTCTCTGG
CCCTCTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCGTCTCATTGGGGAGAAA^{CT}GTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTC
CAGCCTTATATGCAGGAGGTGGTGCCCTTCCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGCTCTGAGAAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCAAAAGCGATTTTTAACCAAAAGGAAGATGGAAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATAATTGATAACATTCAATTGTAACTGGTGTTC
TATACACAGAAAACAATTATTAAATTGCTTTCCATAAAAAAGATTACTTTCCAT
TCCTTAGGGGAAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTATATTATAAA
TGTATTATTATTATTATAAGACTGCATTTATTATATCATTATTATAATATGGATTATTAT
AGAAACATCATTGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKVSSFLMGTLATSCLLLALLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFQPSDFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKGESGEIKAIGELLLLMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CACTCATGGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTGGGCTGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTATTAGCCTTTCCTACAGGTGGTTGCAT
TCTTGCAATGGTCATGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTAACACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTGAACCGGCTCCCCCAGGACCTGTACCACGCCGT
TGCCTGTGCCCGACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTAGCTTGTGTGTGCGGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGGCCAGGTGTACA
ACCACTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGCTTGGGAAAACCTGCACTCTGCACATTTGAAAAGAG
CAGCTGCTGCTAGGGCCGCCGAAGCTGGTGTCTGTCACTTCTCTCAGGAAAGGTTTCAA
GTTCTGCCATTTCTGGAGGCCACCACTCCTGTCTTCCCTTCCATCCCTGCTACCTG
GCCAGCACAGGCACCTTCTAGATATTCCTGCTGGAGAAGAAAGGCCCTGGTTTATT
TGTTGTTACTCATCACTCAGTGAGCATCTACTTGGTGCATTCTAGTGTAGTTACTAGCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTATCAAATAATAT
CTTTATTAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHTYSHWPSCCPKGQDTSEELLRWSTVPVPPLEPA
RPNRHPECRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**A**TGTGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCGAGAGCC
GACC GTTCAATGTGGCTCTGAAACTGGGCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGGTCAACAGGGACTATTCA
ATTTGATGAATGTAAGCTGGTACTCCGGGAGATGCCAGCATCCGTTGTAAGGCCACCAA
GATTTGTGTGACGGGAAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGGG
CCTTCCAGACTCAGACCAGACCCCTGGTGTAAATGGACATTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTCATTGGGCCATAATATTCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTACCTCACCAAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAGGCCGGAAAGCCTGTGGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTACAACCACTCCCCTGGAAACAGATACTGGCTTTATCCAACACAGCAC
TATCATCGGGTTTCTCAGGTGTTGAGCCACACCAGAAGAAACAAACGCGAGCTCAGTGGTGA
TTCCAGTGA~~C~~TGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTCTACTTGTGGC
AGCGACTGCATCCGACATAAGGAACAGTTGTGCTGCCCACAAACAGGCCTTCCCT
GGATAACAACAAAAGCAAGCCGGAGGCTGGCTGCCTCCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTCCTT
TCTACCACCAACTACTGCCCCCATTAAAGGTTCTTGTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTCAAAACATTGCAGAAGTGAGGTCACTCTGAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGTCCAGTGCAGTGGCTTGCACCAAAGAAGGCA
GCAGACAAAGTCGTTCCCTTCCAATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAAGAACTCTCAAGACCTCTCCCTGCCTTAAACCTTCTGCA
GTGATCTAAGAAGCCAGATTCTGCACAAATACGTGGTGGTCACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTAGTGTCTGCCCAAGTACCACTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTTCTCCATGTCAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGT**AG**

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNVS梧
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEGDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCQTGVFPFLDNNK
SKPGGWLPLLLSSLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRCSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-492

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCC~~T~~GATGCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGCTTGCCTTCTGAGTGAGCAGCTCGAAAATCCCCAAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTCCATGTACGTAACATCGAGAGCCGCTCACCTCCCCCTG
GAATTACACTGTCACTGGGACCCCAACGGTACCCCTCGGAAGTTGTACAGGCCAGTGAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
GAGACCC~~T~~GGTGTCCGGAGGAAGCACCAAGGCTGCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGA~~C~~TTGGCTGCACCTGCGTCACCC~~T~~GTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPAMVKYLLSILGLAFLSEAARKIPKGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRLGCINAQGEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCAGG
ACAGAGAGTGACAAAATACCCAGCACAGCCCCCTCCGCCCCCTCGGAGGCTGAAGAGGGATT
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGGCAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTTCTGCTGCCT
TGGCACTGGGCCGAAGCCCAGTGGTCTTCTCTGGAGAGGCTGTGGGGCCTCAGGACGCTACC
CACTGCTCTCGGCCCTCTCGCCGCTCTGGACAGTACGACATACTCTGCCCTGCCCTGGGACAT
CGTGCCCTGCTCCGGGCCGCTGGCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCGTGACATGGGACTGG
GAAGAGCCTGAAGATGAGGAAAGTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCCTCTCTCCAGGCCAAGTCGTGCTCTCCAGGCCTACCCACTGCCCCGTGCGTCTGC
TGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAAGTTGGTCAGTCTGTGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTTAACTCAGCCAGGTACGAGAA
GGAACATCAACCACACACAGCAGCTGCCCTGCCCTGGCTCAACGTGTGAGCAGATGGTACA
ACGTGCATCTGGTTCTGAATGTCTTGAGGAGCAGCAGCTCGGCCCTCCCTGACTGGAATCAG
GTCCAGGGCCCCC AAAACCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTGAA
CCACACAGACCTGGTCCCTGCCCTGTATTCAAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCCCTGAGAGCTGGCTGCTGGACGCACCGTGCCTGCCCGCAGAACGGCACT
GTGCTGGCGGGCTCGGGTGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCTGGAGAACG
TCACTGTGACAAGGTCTCGAGTCCCATTGCTGAAAGGCCACCTAACCTCTGTGTTAGGTG
AACAGCTGGAGAGCTGCAGCTGCAGGAGTGTGACTCCCTGGGCTGACTCCCTGGGCTCTAAAGA
CGATGTGCTACTGTGGAGACAGCAGGGCCCCAGGACAACAGATCCCTCTGTGCTTGGAACCCA
GTGGCTGTACTTCACCTACCCAGCAAAGCCTCCACGAGGGCAGCTGCCCTGGAGAGTACTTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCTG
CCCCATGGACAAATACATCCACAAAGCGCTGGGCCCCCTCGTGTGGCTGGCTGCCACTCTTGCCG
CTGCGCTTCCCTCATCCTCCTCTCAAAAGGATCACCGAAGGGTGGCTGAGGCTCTTGAAA
CAGGACGTCCGCTCGGGGGCGCCGCAAGGGCCCGGGCTCTGCTCCCTACTCAGCCGATGA
CTCGGGTTTCGAGCGCTGGGGCGCCCTGGCGTGGCCCTGTGCCAGCTGCCGTGCGCGTGG
CCGTAGACCTGTGGAGCCGTGGTAAGTGGAGCGCAGGGGGCCCGTGGCTTGGGTTACCGCGCAG
CGGCAGCAGACCCCTGCAAGGAGGGCGCGTGGTGGCTGCTCTCCCGGTGCGGTGGCGCT
GTGCAAGCGAGTGGTACAGGATGGGTGTCCGGGCCCCGGCAAGGGCCGCAAGACGCCCTTCC
GCCCTCGCTCAGCTGCGTGTGCCGACTTCTGCAAGGGCCGGCCGGCAGCTACGTGGGG
GCCCTCGCTCAGGGCTGCTCCACCCGGACGCCGTACCGCCCTTTCCGCACCGTGCCGTCTT
CACACTGCCCTCCCAACTGCCAGACTTCTGGGGCCCTGCAAGCAGCCTGCCGTGCCGTCTTCC
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCGGGCCCTCAGCAGCCCTGGATAGCTACTTC
CATCCCCGGGACTCCCGCCGGACGCCGGGGTGGGACCGAGGGCGGGACCTGGGGCGGGGA
CGGGACTTAATAAAGGCAGACGCTGTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSSLERLVLGPQDATHCSPGLSCRILWSDSILCLPGDIVPAPGPVLAPTHLQTELVLRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLFSQAYPTARCVLLEVQVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKEELNHTQQLPALPWLNVSADGDNVHLVLNVEEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDHSVRTNICPFREDPRAHQNLWQAARLRLLTLQSWLLDAPCSLPAEAALCWRAPIGGDPCQPLVPPLSWENVTVDKLEFPLLKGHPNLCVQVNSSEKQLQLQECLWADSLGPLKDVLLETRGPQDNRSILCALEPSGCTSPLSKASTRAARLGEYLLQDLQSQCQLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAAALSLLLLLKKDHAKGWRLLKQDVRSGAAARGRAALLLYSADDSCFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRQTLQEGGVVLLFSPGAVALCSEWLQDGVSAGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTPVFTLPSQLPDFL GALQQPRAPRSGRLQERAEVQSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251, 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617, 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGCCAGCCCCGATGAGGACGCTGCTGACCATCTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCAGCAACTTTGA
AAACATCCTGACGTGGGACAGCGGGCAGAGGGCACCCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGAAAGAAGGGTGTCAAGGGATCACCCGGAAAGTCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGTCACCGCT
GTCAGTGCGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTCACTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATGATTGAGATGATTGTTCATCTA
CCCCCACGCCAACCCGTGCAGCGATGGCACCGGCTAACCTGGAAGACATCTTCCATGACCTG
TTCTTACCAACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAAGCAGAGA
ATATGAGTTCTTCCGGCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTGCGTTCCA
CCTGGGCCAAGGAGAGTGGCCCTACATGTGCCAGTGAGAAGACACTGCCAGACCGGACATGGACC
TACTCCCTCTCCGGAGCCTCTGTCTCCATGGGCTTCCCTGTCAGTACTCTGCTACCTGAG
CTACAGATATGTCAACCAAGGCCCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
TCCAGCGCTGCCTCATCCAGGAGCACGTCTGATCCCTGTCTTGAACCTCAGCGGGCCCGAGC
AGTCTGGCCAGCCTGTCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGGCCGCAGGAGC
TCCACAGCGCATAGCCTGTCAGTACCTACTTAAAGGCAGCCAGACATCTCATCTCCAGC
CCTCAACAGTGCACCTCCCAAGATCCTCTCCCAACTGTCCTATGCCCAAAACGCTGCCCTGAG
GTCGGCCCCATCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCATTACGCCCCACA
GGCCATCTCAAGGTCCAGCCTCCTCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGTATGCAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTCTAGTCCTAAA
CACCTTAAGGCCCTAAAGGTCAAGCTCAGAAAGAGCCACCAGCTGGAAGCTGCATTTAGGTGGCCT
TTCTCTGCAGGAGGTGACCTCTTGGCTATGGAGGAATCCAAGAAGCAAATATTGCAACAGC
CCCTGGGATTGCAACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGACA
CCACAGTACCTAAAGGGCCAGCTCCCCCTCTCTCAGTCCAGATCGAGGGCCACCCATGTC
CCTCCCCCTTGCAACCTCTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGCC
TGCTGGAGTCCCTGTGTCCAAAGGATGAAAGCCAAGGCCCCCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACTGGATTCTCTTTCAGAGGCCCTGGCCCTGACTGTGCAAGTGGGAGTCTG
AAGGGAAATGGGAAAGGCTTGGCTTCCCTGTCCCTACCCAGTGTCACTCCTGGCTGTCA
ATCCCATGCCCTGCCATGCCACACACTCTGCGATCTGCCCTCAGACGGGTGCCCTTGAGAGAAC
AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCTCACCGAACAAAGCAGCATGATA
AGGACTGCAAGGGGGAGCTCTGGGAGCAGCTGTGAGACAAGCCGTGCTCGTGAGCCCTG
CAAGGCAGAAATGACAGTGCAGGAGGAATGCAGGGAAACTCCCGAGGTCCAAGGCCACCTC
CTAACACCATGGATTCAAAGTCTCAGGGAAATTGCTCTCCCTGCCCTTGTCAATTGTTCAAAGGTGGAAAGAGA
GCCTGGAAAAGAACCAAGGCCCTGGAAAAGAACCAAGCAAGGAGGCTGGCAGAACCAAACACTGC
ACTTCTGCCAAGGCCAGGGCCACAGGACGGCAGGACTCTAGGGAGGGTGTGGCCTGCAGCTCA
TTCCCAAGGCCACTGCCCTGACGTTGACGATTCTAGCTTCAATTCTCTGATAGAACAAAGC
GAATGCAAGGTCCACCAGGGAGGGAGACACACAAGCCTTTCTGCAGGCCAGGAGTTCAAGACCC
ATCTGAGAAATGGGGTTGAAAGGAAGGTGAGGGCTGIGGCCCTGGACGGTACAATAACACAC
TGTACTGATGTCACAACATTGCAAGCTCTGCTTGGGTTGAGCCCATCTGGCTCAAATTCCAGC
CTCACCACTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTCTC
ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTAAAGTGTAAATAGTGCCTGGTACATGGCAGTGCCAATAAACGGTAGCTATTTAAAAA
AAAAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWSGPEGTPDTVYSIEYKTYGERDW
VAKKGCRITRKSCNLTVEGNLTELYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVT CIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFGAFLFSMGFLVAVLCYL SYRYVTKPPAP
PNSLNQQRVLTFQPLRFIQEHVLI PVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILOPSNVPPPQILSPSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLLSPKHLRPKGQLQKEPPAGSCMLGGI SLQEVTS LAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPM SLP LQPPSGPC
SPSDQGPSPWGLLESLVCPKDEAKSPAPETSDLEQPTELDSLFRGLALT VQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCGCGTGGCCACAACATGG
CTGCGCGCCGGGCTGCTCTGGCTGTTGCTGGACCTCAAAGTGTGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGAAAGCTCTGAAGACTTCACGGGCCCTGATTGTCGTTGTGAATTAA
AAAAAGGTGACGATGTATATGTTACTACAAAACTGGCAGGGGATCCCTGAACTTGGCTGGA
AGTGTGAACACAGTTGGATATTTCAAAAGATTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTGTCTGCTTGAAAGGAGGAAGAGATGATT
TTAATAGTTATAATGAGAAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAACAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCCTGTGCCCTGAGCCCGAGGCATTCAGAGCTGATTCAGAGGATGGAGAAGGTGCTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGGCCACCCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTCGTTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCGTTATT
CATTACAGCAAAGGATTTCGTTGGCATAAACTAAGTTGTTACAAAGATTGTTTTAGTA
CTAAGCTGCCCTGGCAGTTGCATTTTGAGCCAACAAAAATATATTTTCCCTCTAAGTA
AAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVCDEECMILMYRGKALEDFTGPDCRFVN
FKKGDDVVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPADETDFVCFEGGRD
DFNSYNVEELLGSLELEDSPPEESKKAAEVSQHREKSPPEESRGRELDPVPEPEAFRADSEDGEGA
FSESTEGLQGQPSAQESHPHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCIVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCAAAGCCGGGCTTCTGCCTCGCCTCTAGGACATACACGGGACCCCCCTAACCTC
AGTCCCCAAACGCGCACCCCTGAAGTCTTGAACCTCAGCCCCGACATCCACCGCGGGCACAGG
CGCGGCAGGCAGGCGAGGTCCCAGCGAAGGGCATGCGCAGGGGGTCGGGCAGCTGGCTCGGGC
GGCGGGAGTAGGGCCCGCAGGGAGGCAGGGAGGCTGCATATTCAAGAGTCGCAGGGCTGCCTG
GGCAGAGGCCGCCCCCTCGCTCCAGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGGTGG
TCTCGCTGCTGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTCTGCCGCCGCGTGGTCAGC
GCCAAAAGGTGTGTTTGCTGACTTCAGCAGTCAAGCATTGCTACAAAATGGCTACTTCCATGAAC
GTCCAGCCGAGTGAGCTTCAGGAGGCAGCCCTGGCTGTGAGAGTGAGGGAGGAGTCCTCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGG
ACAGGGATTTCTGATGGTGATTCTGGATAGGGCTTGGAGGAATGGAGATGGCAAACATCTGG
TGCCTGCCAGATCTTACCACTGGTCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTTCCCTGCGGAAGTAAAAGTGTGTTGATGATCAGGAACTGCCAATCTGG
CTTGGGGTCCCTACCTTACCACTGGAAATGATGACAGGTGTAACATGAAGCACAATTATTTG
CAAGTATGAAACAGAGATTAATCCAACAGCCCCCTGTTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCATGAAATGTGGTTACTGAAGCAGGTATAATTCCAATCTAATTATGGTT
ATACCAACATAACCCCTGCTCTTACTGATACTGGTTGTTGGAACCTGTTGTTCCAGATGCT
GCATAAAAGTAAGGAAGAACAAAAACTAGTCAAACCCAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTAATAACTCATTGACTTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTTGAATGGCTGAAATCACAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTGAGGCAAATATTAAAGTAATTGTTATGCTATTATTC
TTTAAAGAATATGCTGTGCTAATAATGGAGTGGACATGCTTATTGCTAAAGGATGCCAA
ACTTCAACACTCAAGAACATGAAATGGACAATGCAGATAAAGTGTATCAACACGTCGGAGTA
TGTGTGTTAGAACAACTCCCTTATTCTTACCTTCACTGTTGCTAAAGTGTATCTAGTCAGTAA
TGTATATTGTTGAAATTTCAGTGTGCAAAGTATTGTTACCTTGCATAAGTGTGTTGATAAAA
ATGAACCTGTTCTAATATTGTTATGGCATCTCATTTCAATAACATGCTTTGATTAAAG
AAACTTATTACTGTTGCAACTGAATTCAACACACACAAATATAAGTACCATAGAAAAAGTTGT
TTTCTCGAAATAATTCACTCTTCAGCTCTGCTTGGTCAATGCTTAGGAAATCTCTCAGA
AATAAGAAGCTATTCAATTAGTGTGATATAACCTCTCAAACATTACTAGGAGGCAAGGAT
TGTCTAATTTCATTGCAAGACATGTGCTTATAATTATTAGCTTAAATTAAACAGATT
TTGTAATAATGTAACTTGTTAATAGGTGCTAAACACTAATGCAGTCATTGAAACAAAAGAAG
TGACATACACAATATAAAATCATATGCTTACACGTTGCCTATATAATGAGAACAGCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTTGCCCACTAAACAAAGATGGTTGTTGGGGTTGG
ATTGACACTGGAGGCGAGATAGTTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTTAGCTCTG
ACTATATTAGTATACAAAGAGGTATGTGGTTGAGACAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTAGGAAGGAAAGGAACATCGAAATCGTGTGAAATGGGGTTGG
AACCCATCAGTGATCGCATATTCAATTGATGAGGGTTGCTTGAAGATAGAAAATGGTGGCTCTT
CTGCTTATCTCCTAGTTCTCAATGCTTACGCCCTGGCTCTTAAACAAATAAGAGTCTTGTGTTCTGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSSLGAALLCGHGAFCRRVSGQKVCFADFKHPCYKMAFHESRVSFQEALACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217